Measures of 3D medical image alignment.

Studholme, Colin

Measures of 3D Medical Image Alignment

Colin Studholme

A Thesis Submitted for the Degree of Doctor of Philosophy
of the University of London

August 1997

Computational Imaging Science Group,
Division of Radiological Sciences,
United Medical and Dental Schools of Guy's and St Thomas's Hospitals

Abstract

The work covered in this thesis deals with the problem of automatically aligning 3D images acquired from different medical imaging modalities. The approach taken is to develop generic measures of image alignment derived from the co-occurrence of values in the two images. The development of statistical alignment measures is reviewed. The alignment problem is then expressed in terms of entropy and developed using tools from information theory. The problem of normalisation with respect to image overlap is identified as important in this application. Normalised mutual information is proposed and compared to other information theoretic measures of alignment.

The key factors in the design of a system which efficiently evaluates and optimises similarity between 3D volumetric image values are described. Experiments are carried out using this system to compare the behaviour of 6 different measures in recovering the alignment of MR-CT and MR-PET brain images. These confirm that normalised mutual information provides the best overall measure of alignment. The performance of an automated registration system using this measure is then validated over a large range of clinical data.

An extension of entropy derived measures of alignment is then proposed to enable the automated alignment of images containing severe intensity distortion. This is achieved by using an extra channel of information to evaluate alignment across a set of image partitions. Experimental results are included for MR-PET registration in the pelvis and surface coil MR to CT registration in the brain.

Finally there is discussion of the limitations of intensity based similarity as a measure of alignment. An approach is developed to introduce knowledge of region topology into the alignment measure. Results are included illustrating how this may provide a useful measure of alignment for some cases of MR and PET images of the pelvis, where conventional intensity based measures fail.

Acknowledgements

I am particularly grateful to my supervisor Dr D.J. Hawkes whose experience in medical imaging and registration provided the foundation for this work. In addition I must thank my other main proof reader and source of mathematical inspiration Dr John Little. I also must not forget all the other people who have contributed directly or indirectly to my ideas on this work over the last four years. These include, in alphabetical order (without titles!):

Dave Atkinson, Ramsey Badawi, Charlie Bird, Alan Colchester, Tim Cox, Iain Cranston, Eddie Edwards, Colin Ferrie, Frans Gerritsen, Mike Gleeson, Lewis Griffin, Derek Hill, Steve Kevil, Mike Kopelman, Pete Liepins, Mandy Lewis-Gormen, Petra Lewis, John Lynch, Mike Maisey, Nimrat Manget, Paul Marsden, Greame Penny, Colin Renshaw, Glynn Robinson, Cliff Ruff, Bryan Scattergood, Karel Strasters, Tony Strong, John Studholme, Paul Summers, Joseph Wong, Michael Wong, Wai-Lup Wong, Jason Zhao.

The work in this thesis would not have been possible without funding from the U.K. EPSRC, Philips Medical Systems (ICS/AD) Best, and the PET centre of Guy's and St Thomas' Hospitals.

The work has benefited greatly from access to a wide range of clinical data for which I acknowledge the radiographic staff of Guy's and St Thomas's (The PET centre, MR centre and CT), St Georges (MR unit) and the Maudsley (CT) Hospitals London. The images with fiducial marker standard transformations used in chapter 6 and part of chapters 7,8 and 9 were provided as part of the project "Evaluation of Retrospective Image Registration", National Institutes of Health, Project Number 1 R01 NS33926-01, Principle Investigator, J. Michael Fitzpatrick, Vanderbilt University, Nashville, TN. In addition, for the use of the Covira MR-CT data in chapter 7, the radiologists W.P.Th.M. Mali MD, PhD and L. Ramos MD and neurosurgeon C.W.M. van Veelen MD, PhD from Utrecht University Hospital (AZU). Finally for one of the MR-PET image pairs used in chapter 8, I thank Ralph Myers, PhD at the MRC Cyclotron Unit, Hammersmith Hospital, London.

Contents

Abstract 3
Acknowledgements 5
Table of Contents 7
List of Tables 13
List of Figures 15
List of Symbols 19
List of Abbreviations 21

1 Introduction **23**
1.1 The Clinical Requirement 23
1.2 Gaining Information by Registration 24
1.3 Hypothesis 28
1.4 An Overview of the Thesis 28

2 Identification and Alignment of Corresponding Features **31**
2.1 Introduction 31
2.2 The Spatial Relationship between Acquisitions 32
2.3 Corresponding Point Alignment 33
2.3.1 Landmark Identification 33
2.4 Corresponding Surface Alignment 34
2.4.1 A Distance Between Corresponding Surfaces 34
2.4.2 Early Approaches 36
2.4.3 Surface Extraction and Correspondence 36
2.4.4 Outliers and Surface Extent 41
2.4.5 Using Related Surfaces 43
2.5 Combining Different Types of Geometrical Features 43
2.5.1 Sequential Feature Alignment 44
2.5.2 Simultaneous Feature Alignment 44
2.6 Generic Feature Alignment 45
2.7 Voxel Based Object Alignment 45
2.8 Summary 46

3 Image Similarity As a Measure of Alignment **49**
3.1 Introduction 49
3.2 Views of Anatomy and Physiology 50
3.2.1 Materials, Properties and the Delineation of Regions 50
3.2.2 Combining Views of a Scene 52
3.2.3 Spatial Correspondence 53
3.2.4 Registration Criteria 53
3.2.5 Statistics: A Global View of Region Alignment 57
3.3 Early Approaches 58
3.3.1 Aligning Images of the Same View 58
3.3.2 Enforcing Similarity:Correlation of Pre-Processed Images 61
3.4 Statistical Measures between Different Scenes 63
3.4.1 Corresponding Intensity Variance 64
3.4.2 Moments of the Joint Probability Distribution 65
3.4.3 Clustering and Labelling 65
3.4.4 Describing and Optimising the Intensity Transformation 66
3.5 Summary 67

4 Information Theory and Image Alignment **69**
4.1 Introduction 69
4.2 Information and Entropy 70
4.2.1 Information from Uncertainty 70
4.2.2 A Measure of Information 71
4.2.3 Information Measures from Two Views of a Scene 72
4.3 Information Measures and Alignment 73
4.3.1 Joint Entropy and Image Alignment 73
4.3.2 Normalisation: The Need for Relative Entropies 74
4.3.3 Mutual Information 75
4.3.4 Further Constraints 77
4.4 The Image Overlap Problem: A Simulation 79
4.4.1 Image Overlap: A Simple Model 79
4.4.2 Response of Entropy Measures to Varying Field of View 82
4.5 Summary 86

5 Evaluation and Optimisation of Image Similarity **87**
5.1 Introduction 87
5.2 From Properties to Measurements: Noise and Spatial Resolution 88
5.2.1 Measurement Noise 88
5.2.2 Measurement Resolution 94
5.3 Evaluating Spatially Corresponding Measurements 95
5.3.1 An Engineering Compromise: Two Stage Resampling 96

5.3.2 Isotropic Resolution 98
5.3.3 Interpolation During Iterative Alignment 100
5.4 Forming an Estimate of the Joint Probability Distribution 101
5.4.1 Discrete versus Continuous 101
5.4.2 Binning Values 102
5.5 Alignment Search Space 103
5.5.1 Multi-Resolution Alignment 105
5.6 Multi-Resolution Optimisation 106
5.6.1 Relating Changes in Translation and Rotation 107
5.6.2 Initial and Final Step Size 108
5.7 Summary 109

6 An Experimental Comparison of Alignment Measures 111
6.1 Introduction 111
6.2 MR-CT and MR-PET Brain Image Alignment 112
6.2.1 Distributions of MR and CT Image Values 112
6.2.2 Distributions of MR and PET Image Values 115
6.2.3 Measures 116
6.3 The Vanderbilt Image Registration Database 119
6.3.1 An Independent Registration Estimate 119
6.3.2 Image Data 120
6.4 Assessing Registration Quality 120
6.4.1 Expressing Accuracy Relative to a Gold Standard 120
6.4.2 Visual Evaluation of Image Alignment 122
6.5 Robustness and Accuracy of Registration with Respect to Image Content 127
6.5.1 Experimental Design 127
6.5.2 Results: MR-CT 127
6.5.3 Results: MR-PET 132
6.6 Capture Range and Imaged Field of View 135
6.6.1 Exploration of Alignment Parameter Space 135
6.6.2 Experimental Design 137
6.6.3 Analysis of Results 138
6.6.4 Results: MR-CT 139
6.6.5 Results: MR-PET 141
6.7 Summary 142

7 Retrospective Alignment of Clinical MR and CT of the Head 143
7.1 Introduction 143
7.2 The Effect of Image Noise and Resolution on Alignment Accuracy 145
7.2.1 Experimental Method 145
7.2.2 Results: Image Resolution 148

7.2.3 Results: Measurement Noise . . . 149
7.2.4 Discussion . . . 149
7.3 A Database of Clinical MR-CT Images . . . 151
7.3.1 Data Acquired Locally . . . 151
7.3.2 Image Data from the COVIRA Database . . . 152
7.3.3 A Summary of Clinical Image Characteristics . . . 153
7.3.4 Independent Manual Estimate of Registration . . . 157
7.3.5 Common Types of Clinical Image Misalignment . . . 159
7.4 Automated Rigid Registration . . . 163
7.4.1 Algorithm Parameters . . . 163
7.4.2 Experimental Design . . . 163
7.4.3 Accuracy . . . 164
7.4.4 Precision . . . 165
7.5 Recovery of Scaling and Skew Parameters . . . 168
7.5.1 Why Include Scaling and Skew Estimates? . . . 168
7.5.2 Incorporating Scaling and Skew in the Transformation Matrix . . . 168
7.5.3 Optimisation of Scale and Skew Parameters . . . 169
7.6 Scaling and Skew Recovery from Clinical Data . . . 170
7.6.1 Experimental Design . . . 170
7.6.2 Results . . . 170
7.7 Summary . . . 177

8 Retrospective Alignment of Clinical MR and PET of the Brain 179
8.1 Introduction . . . 179
8.2 The Effect of Image Noise and Resolution on Alignment Accuracy . . . 180
8.2.1 Experimental Method . . . 180
8.2.2 Results: Image Resolution . . . 183
8.2.3 Results: Measurement Noise . . . 184
8.2.4 Discussion . . . 184
8.3 A Database of Clinical MR-PET Images . . . 186
8.3.1 An Overview of Clinical Image Characteristics . . . 186
8.3.2 An Independent Manual Estimate of Registration . . . 191
8.3.3 Common Types of Clinical Image Misalignment . . . 191
8.4 Automated Rigid Registration . . . 192
8.4.1 Algorithm Parameters . . . 192
8.4.2 Experimental Design . . . 192
8.4.3 Results . . . 193
8.5 Summary . . . 198

9 The Alignment of Images With Severe Intensity Distortion **199**
9.1 Introduction . . . 199
9.2 Measurement Inhomogeneity . . . 200
9.2.1 Measurement Similarity and Inhomogeneity . . . 200
9.2.2 A Solution: Image Partitioning . . . 201
9.3 Encoding Spatial Location . . . 201
9.4 Aligning Images with Known Distortion Geometry . . . 203
9.4.1 Introduction . . . 203
9.4.2 Aligning MR and PET of the Pelvis . . . 203
9.4.3 MR intensity Inhomogeneity . . . 205
9.4.4 Evaluation and Optimisation of Image Similarity . . . 207
9.4.5 Experimental Design . . . 208
9.4.6 Results . . . 209
9.4.7 Discussion . . . 210
9.5 Automated Partitioning to Recover Unknown Distortion Geometry . . . 215
9.5.1 A Measure of Partitioning . . . 215
9.5.2 Partitioning and Alignment . . . 217
9.6 Automated Alignment of MR Surface Coil Angiography of the Brain . . . 219
9.6.1 Surface Coil MR Imaging . . . 219
9.6.2 Alignment of MR Angiography . . . 219
9.6.3 Evaluation and Optimisation of $I(M, \mathcal{X}; N)$ . . . 219
9.6.4 Experimental Design . . . 221
9.6.5 Results . . . 223
9.7 Summary . . . 228

10 Region Topology and Image Similarity **229**
10.1 Introduction . . . 229
10.2 When Voxel Similarity Fails: Why we still Need High Level Vision . . . 230
10.3 Using Region Topology in Image Registration . . . 231
10.3.1 Connected Component Labelling . . . 231
10.3.2 Using Connected Component Information in Image Alignment . . . 231
10.3.3 Information and Region Labelling . . . 235
10.4 The Application of Region Labelling to MR-PET Pelvis Image Alignment 238
10.4.1 Method . . . 238
10.4.2 Results . . . 239
10.5 Summary . . . 241

11 Conclusions **243**
11.1 Summary of Findings . . . 243
11.1.1 Similarity Measures . . . 243
11.1.2 Implementation . . . 243

11.1.3 Comparison of Similarity Measures 244
11.1.4 Automated Brain Image Registration 244
11.1.5 Extension of Voxel Similarity Techniques 244
11.2 The Clinical Significance 245
11.2.1 Robustness and Automated Registration 245
11.2.2 Accuracy and Uncertainty 245
11.3 Further Work 247
11.3.1 Extension of Transformation Complexity 247
11.3.2 Similarity Measures 247
11.3.3 Increasing the Speed of Registration 248
11.3.4 Intra-operative Imaging 249

Bibliography **251**

Selected Publications **261**

List of Tables

6.1 MR-CT: Successful Registrations as Transaxial Field of View is Varied . . 140
6.2 MR-CT: Successful Registrations as Axial Field of View is Varied 141
6.3 MR-PET: Successful Registrations as Transaxial Field of View is Varied . 142

7.1 MR Image Volume Parameters 152
7.2 CT Image Volume Parameters 153
7.3 Manual MR-CT Registration Estimates 161

8.1 MR Image Volume Parameters 189
8.2 Manual MR-PET Registration Estimates 191

9.1 MR(Head Coil)-MR(Surface Coil): Registration Estimate Parameters . . 226
9.2 MR(Surface Coil)-CT: Registration Estimate Parameters 226

10.1 Table of Rigid Parameters Derived With and Without Labelling 240

List of Figures

1.1 Example of the Structures Delineated by MR and CT of the Head 25
1.2 Example of the Structures Delineated by MR and CT of the Head 26
1.3 Example of the Structures Delineated by MR and PET Scans of the Brain 26
1.4 Example of the Structures Delineated by MR and PET Scans of the Pelvis 27

2.1 A Measure of Alignment Between Corresponding Surfaces 34
2.2 Points Describing Corresponding Outlines 35
2.3 Distance Between Related Boundaries 39
2.4 The Closest Point On a Surface 40
2.5 Limited Volume of Overlap Influencing Surface Distance 41
2.6 Outliers Caused by Difference in Structure 42

3.1 Alternative Views Provided by Different Modalities 52
3.2 A Map of Properties Delineating Regions of Material 53
3.3 A Combined set of Regions 54
3.4 Volume of Overlapping Regions Varying with Transformation 55
3.5 A More Subtle Point of Alignment 55
3.6 Alignment Constrained by Two Regions 56
3.7 Region and Boundary Correlation 63

4.1 Information Content of a Binary Image 72
4.2 Anatomical Structure and Information 73
4.3 Sketch of Alignment by Removal of Duplicated Features 74
4.4 Venn Diagram of Joint and Marginal Entropies 76
4.5 Additional Constraints on Shared Information 77
4.6 Balancing Unshared Information 78
4.7 Sketch of the Truncation of Overlapping Regions by the Field of View . . 79
4.8 A Simple Model of a Scene to Compare Entropy Measures 80
4.9 Evaluating Area of Overlapping Field of View 80
4.10 Evaluating Overlapping Objects in Scene 81
4.11 The Change in Response of $-H(\tilde{M};\tilde{N})$ as the Field of View is Varied . . 82
4.12 The Change in Response of $I(\tilde{M};\tilde{N})$ as the Field of View is Varied 83
4.13 The Change in Response of $Y(\tilde{M};\tilde{N})$ as the Field of View is Varied . . . 84

4.14 The Change in Response of $C_e(\tilde{M};\tilde{N})$ as the Field of View is Varied . . . 85
4.15 The Change in Response of $\rho(\tilde{M};\tilde{N})$ as the Field of View is Varied 85

5.1 CT Noise Distribution 89
5.2 CT Noise with Storage and Reconstruction Artefact 90
5.3 MR Noise Distribution 90
5.4 PET Noise Distribution 91
5.5 Example Joint Probability Distribution of MR-CT Values 92
5.6 Example Joint Probability Distribution of MR-PET Values 93
5.7 Image Resolution and the Joint Probability Distribution 95
5.8 Image Resolution and the Response of the Joint Probability Distribution . 96
5.9 The Sampling Rate and Resolution of a 3D Image 97
5.10 Cubic Resampling of 3D Images 98
5.11 Varying Kernel Overlap Threshold 100
5.12 The Range of Values in an Image that we need to Consider 103
5.13 The Effect of Image extent on Alignment Parameter Space 104
5.14 The Capture Range 106
5.15 Relating Changes in Rotations and Translations 107

6.1 Response of MR(T1)-CT Joint Histogram to Misalignment 113
6.2 Response of MR(PD)-CT Joint Histogram to Misalignment 114
6.3 Response of MR(T2)-CT Joint Histogram to Misalignment 115
6.4 Images of MR(T1)-CT Joint Histogram 115
6.5 Response of MR(T1)-PET Joint Histogram to Misalignment 116
6.6 Response of MR(PD)-PET Joint Histogram to Misalignment 117
6.7 Response of MR(T2)-PET Joint Histogram to Misalignment 118
6.8 Images of MR(T1)-PET Joint Histogram 118
6.9 CT Voxels Used to Assess Registration Accuracy 122
6.10 Point Displacement between Two Transformation Estimates 123
6.11 PET Image Voxels used to assess Registration Accuracy 124
6.12 Visualisation of MR-PET Alignment 125
6.13 Visualisation of MR-CT Alignment 126
6.14 MIPVD of MR-CT Estimates Using Statistical Measures 128
6.15 Example of a Failed MR-CT Registration using $\sigma(M)$ 129
6.16 MIPVD of MR-CT Estimates Using Entropy Measures 130
6.17 Example of a Failed MR-CT Registration using $\Gamma(M,C)$ 130
6.18 Example of a Failed MR-CT Registration using $H(M,C)$ 131
6.19 Example of a Failed MR-CT Registration using $I(M,C)$ 131
6.20 MIPVD of MR-PET Estimates Derived from Statistical Measures 132
6.21 MIPVD of MR-PET Estimates Using Entropy Measures 133
6.22 Comparison of Fiducial Marker and Automated Estimate 134

6.23 Examples of MR Transaxial Field of View 138
6.24 Examples of CT Transaxial Field of View 138
6.25 Examples of PET Transaxial Field of View 139
6.26 Example of MR and CT Axial Field of View 139

7.1 Examples of MR-CT Image Resolutions 146
7.2 Examples of MR-PET Images with added Gaussian Noise 147
7.3 MIPVD as Sampling Resolution is Varied 148
7.4 Variation of MIPVD as Added Noise Level is Varied 149
7.5 Examples of Range of Axial Extent of CT in Clinical Imaging 154
7.6 Example of Limited in-plane Extent of CT Image Reconstruction 155
7.7 Example of Artefacts due to Metal Components of Stereotactic Frame . . 156
7.8 Example of Artefacts due to Fillings in Teeth 157
7.9 Example of Deformable Tissues Imaged in the Neck 157
7.10 Example of Patient Step Motion Artefact in CT 158
7.11 Example of Patient Motion Artefact in MR 160
7.12 Illustration of Alignment Parameters 162
7.13 MIPVD as the number of intensity bins is varied 163
7.14 MIPVD Between Manual and Automated MR-CT Estimates 164
7.15 Example of Automated Solution Visually Superior to Manual Estimate . . 165
7.16 Example of Automated Solution Visually Superior to Manual Estimate . . 166
7.17 Standard Deviation of Rigid Registration Parameters 167
7.18 Relating Changes in Scale and Skew Parameters 169
7.19 Automated Scaling and Gantry Tilt Estimate 172
7.20 Improvement in Registration by Including Scaling and Skew Estimates . . 173
7.21 Misalignment of Stereotactic Frame 174
7.22 Poor Registration using both 6 and 10 Parameter Estimates 174
7.23 Standard Deviation of Scaling and Gantry Tilt Estimates 175
7.24 Standard Deviation of Rigid Parameters for Estimates including Gantry Tilt and Scale 176

8.1 Examples of MR-PET Image Resolutions 181
8.2 Examples of MR-PET Images with Added Gaussian Noise 182
8.3 Variation of MIPVD as Sampling Resolution is Varied 183
8.4 Variation of MIPVD as Sampling Resolution is Varied 184
8.5 Variation of MIPVD as Added Noise Level is Varied 185
8.6 An Example of MR-PET Images of Normal Anatomy 186
8.7 Example of MR-PET Images of Brain with Atrophy 187
8.8 MR-PET images of Significant Abnormality in the Brain of a Child 188
8.9 Example of MR-PET image of Temozolomide Uptake 190
8.10 MIPVD as the Number of Intensity Bins is Varied 193

8.11 MIPVD Between Manual and Automated MR-PET Estimates 194
8.12 Patient M: Automated Solution Visually Superior to Manual Estimate . . 195
8.13 Patient K: Automated Solution Visually Superior to Manual Estimate . . 196
8.14 Standard Deviation of Rigid Registration Parameters 197

9.1 Venn Diagram of Information from 3 Sources. 202
9.2 Example of PET Transmission and Uncorrected Emission Scans 205
9.3 Sketch of Imaging Planes w.r.t. Coil and Patient 206
9.4 Traces Showing MR Intensity Distortion 207
9.5 Simulated MR Distortion in the Pelvis 209
9.6 Graph of Simulated Distortion Function 210
9.7 Example Clinical MR and PET of the Pelvis 211
9.8 Response of Mutual Information Measures to Simulated Intensity Distortion212
9.9 Example MR-PET Pelvis Registration Result 213
9.10 Response of Mutual Information Measures to Clinical Intensity Distortion 214
9.11 Alignment of Partitioning with Intensity Distortion 216
9.12 Planar Partitioning Determined by Two Variables 217
9.13 Partitioned mutual information evaluated at different partitioning angles . 218
9.14 Example of Surface Coil Angiography Flow and MR Signal Images 220
9.15 Interpolation During Evaluation of Joint Histogram 221
9.16 Example of Registered MR Surface and Head Coil Acquisitions 222
9.17 Graph of Intensities Across Head and Surface Coil Images 223
9.18 Image of Bottle Acquired to Capture map of Surface Coil Distortion . . . 224
9.19 Artificially Distorted MR with Corresponding CT 225
9.20 MR(Head Coil)-MR(Surface Coil): Registration Errors 227
9.21 MR(Surface Coil)-CT: Registration Errors 227

10.1 A Sketch of A Scene where Voxel Similarity Fails 230
10.2 A Second Example of A Scene Where Voxel Similarity Fails 231
10.3 An Example of a Connected Component Labelling 232
10.4 An Illustration of the Effect of Labelling on Alignment of 1D Signals . . . 233
10.5 Illustration of the Effect of Thresholding and Labelling on Entropy 237
10.6 Example Slice illustrating Labelling of MR 239
10.7 Graph Showing Improvement of Response with Added Labels 239

List of Symbols

$\mathbf{x}$	Rectangular coordinates of a location within an imaged scene.
V_M	The set of points making up the volume of material imaged by modality M.
$\mathcal{R}$	The set of regions of different material within a scene being imaged.
$\tilde{m}(\mathbf{x})$	The underlying value exhibited by a material at location $\mathbf{x}$.
$m(\mathbf{x})$	The measured value of a physical property at location $\mathbf{x}$.
$\tilde{M}$,$\tilde{N}$	Sets of values of properties exhibited by materials in a scene.
M,N	Sets of measurement values recorded in an image of a scene.
k	A set of points $k \subset V_M$ all exhibiting the same value of a property.
$\mathcal{M}$,$\mathcal{N}$	Sets of regions within a scene each exhibiting a particular value.
$\nu(k)$	The volume of a region k in a scene.
$p(\tilde{m})$	Probability of a value $\tilde{m}$ occuring within an imaged scene.
$p(m)$	Probability of a measurement value m occuring within an image.
$p(m\|\tilde{m})$	Probability of a measurement value m occuring given it was recorded in a material exhibiting value $\tilde{m}$.
T	A transformation mapping between two 3D images of a scene.
t_x, t_y, t_z	Translations in x, y and z axes.
$\theta_x, \theta_y, \theta_z$	Rotations around x, y and z axes.
$\gamma_x, \gamma_y, \gamma_z$	Orthogonal Scaling in x, y and z axes.
$p(\tilde{m}, \tilde{n})$	Probability of values $\tilde{m}$ and $\tilde{n}$ occuring together at a point within an imaged scene.
$p(m, n)$	Probability of measured values m and n being recorded together at a point within an imaged scene.
$\Gamma(M, N)$	Correlation Coefficient between sets of values M and N.
$\sigma_M(N)$	Normalised standard deviation of corresponding image values N with respect to the image values M.
$H(M)$	Entropy of a set of symbols $m \in M$.
$H(M, N)$	Joint Entropy between sets of measured values M and N.
$I(M; N)$	Mutual information between sets of measured values M and N.
$Y(M; N)$	Normalised Mutual information between sets of measured values M and N.
$\mathcal{X}$	A set of image partitions, each consisting of a set of points in the image.
α, β	Angles determining orientation of planar partitioning of an image space.

List of Abbreviations

3D	Three Dimensional.
2D	Two Dimensional.
ADC	Analogue to Digital Converter
CPU	Central Processing Unit
CSF	Cerebrospinal Fluid
CT	(X-ray) Computed Tomography
FDG	Fluorodeoxyglucose
FFT	Fast Fourier Transform
FOV	Field of View
FWHM	Full Width at Half Maximum
HMPAO	Hexamethylpropylene amine oxime: Radiopharmaceutical used in SPECT imaging.
MIPVD	Mean Imaged Patient Voxel Displacement: A measure of the registration accuracy.
MR	Magnetic Resonance
MRA	Magnetic Resonance Angiography
PET	Positron Emission Tomography
RAM	Random Access Memory
RF	Radio Frequency
RMS	Root Mean Square
SAVD	Sum of Absolute Value Difference
SPECT	Single Photon Emission Computed Tomography

Chapter 1

Introduction

The work described in this thesis is concerned with the problem of automatically aligning 3D medical images. In particular it deals with the task of aligning images which contain both corresponding and complementary structure, commonly referred to as the problem of *multi-modality* image registration. The aim is to be able to recover the alignment of images after they have been acquired using conventional imaging protocols, i.e. automated *retrospective* registration. It concentrates on the tasks of defining, evaluating and optimising a generic measure of alignment between different 3D medical modalities.

1.1 The Clinical Requirement

There is an ever increasing range of clinical imaging modalities available, each measuring distinctly different properties or characteristics of material within the patient. This thesis is mainly concerned with the alignment of 3D volumetric images of the head that are acquired for a number of clinical applications. Some work is also included on the registration of images of the pelvis. The main 3D imaging modalities dealt with in this thesis are:

- Computed Tomography (CT) X-ray Imaging.
- Positron Emission Tomography (PET) Imaging.
- Nuclear Magnetic Resonance (NMR or MR) Imaging.

Each of these modalities can be tuned to distinguish significantly different properties within the body by varying acquisition parameters and imaging protocol for example:

- X-ray Energy.
- Chemical Tracer and Isotope.

- MR Enhancement Chemical (for example Gadolinium).
- MR Pulse Sequences.

Different settings of these parameters may provide significantly different contrast between tissues and so produce a distinctly different 'modality'. This is particularly true for MR images where modification of echo times and method of data acquisition can produce images delineating white matter and grey matter, cerebrospinal fluid, blood flow or even brain function. An important limitation is that even these images cannot in general be acquired simultaneously, and therefore, because of unknown patient movements between scans, cannot be assumed to be spatially aligned.

1.2 Gaining Information by Registration

Often in clinical practice, a single modality alone does not provide adequate information about a patient's condition, and so they are imaged by a second (and sometimes a third) modality. The final images are conventionally presented separately to the clinician on printed films (for example MR and CT) or on a computer display or colour printout (SPECT and PET). MR provides high contrast between many different types of soft tissues. CT provides soft tissue as well as high resolution bone structure. With the aid of contrast agents MR (and also CT) is capable of delineating regions of different physiology. PET and SPECT provide very specific and sensitive 'functional' or physiological images, but often with little anatomical detail and low spatial resolution. It is natural to assume that bringing these different sources of information into accurate spatial alignment will provide significant extra clinical information.

The application of manual and semi-automatic registration methods to a number of clinical imaging areas has confirmed that accurate spatial alignment does provide additional clinically useful information. Some of the applications include:

- MR and CT of the head for surgery and radiotherapy planning [47, 54, 56].
- MR and SPECT or PET of the brain to anatomically localise tracer uptake indicating brain physiology [50, 58].
- MR or CT and PET or SPECT of the pelvis and abdomen to localise tracer uptake for cancer detection and staging [83, 106].

Examples of the kinds of images acquired clinically and registered manually at our site for skull base surgery planning are shown in Figures 1.1 and 1.2. These illustrate the power of using MR to distinguish different tissue types, combined with CT which

provides high contrast and resolution between bone and air structures. Knowledge of bone structures in the sinuses is often important in the planning and guidance of a surgical approach to a lesion in the skull base. In Figure 1.1, Gadolinium enhanced MR delineates a small acoustic neuroma, while CT delineates air and bone structure in the temporal sinus which is useful in planning the removal of the tumour. Relating enhancing areas of lesion in MR to the fine bone surfaces in CT, as shown in Figure 1.2, can also be important in determining the extent of lesion growth into the sinuses.

Figure 1.1: Corresponding transaxial slices from manually registered images showing different structures delineated by Gadolinium enhanced T1 weighted MR (left) and CT (right) scans of the head.

Figure 1.3 illustrates the use of MR and PET brain image registration. Here the high resolution MR can provide anatomical information to aid in the interpretation of the tracer uptake in PET. This is of added importance in this example where the underlying anatomy is significantly abnormal due to the presence of a large space occupying lesion.

Figure 1.4 illustrates the use of registered MR and PET images of the pelvis. Here the high resolution anatomical information provided by MR scans of the pelvis provide valuable context for the sparse tracer uptake map provided by PET in this region of the body. This is particularly valuable in the detection and staging of cervical cancer [106]. Using additional $^{18}F^{-}$ tracer in a clinical ^{18}FDG PET acquisition provides enough shared bone structure to allow alignment of the PET image with MR. This approach to registration is discussed in more detail in section 9.4.

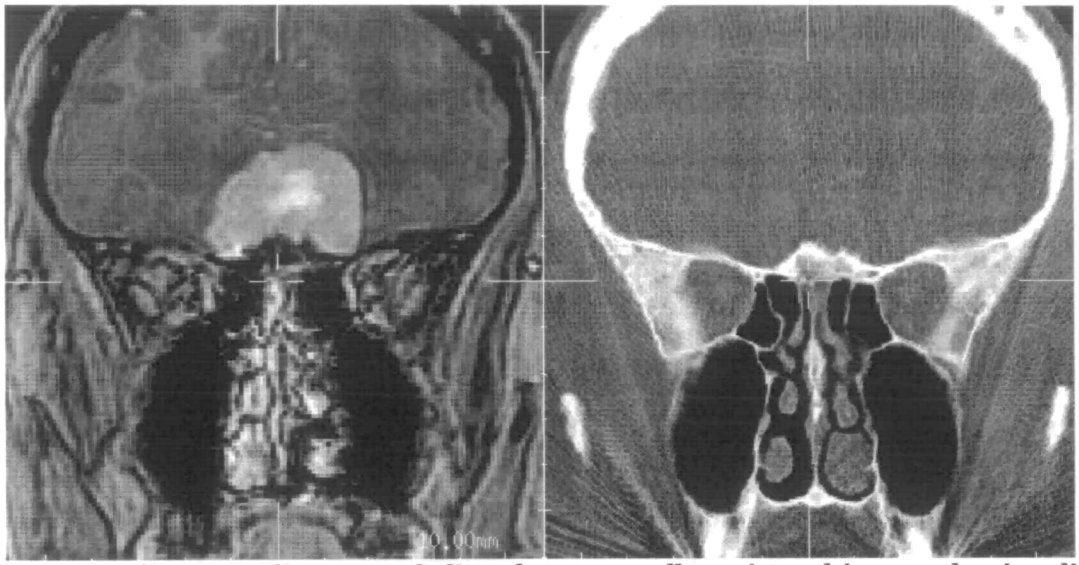


Figure 1.2: Corresponding coronal slices from manually registered images showing different structures delineated by Gadolinium enhanced T1 weighted MR (left) and CT (right) scans of the head.

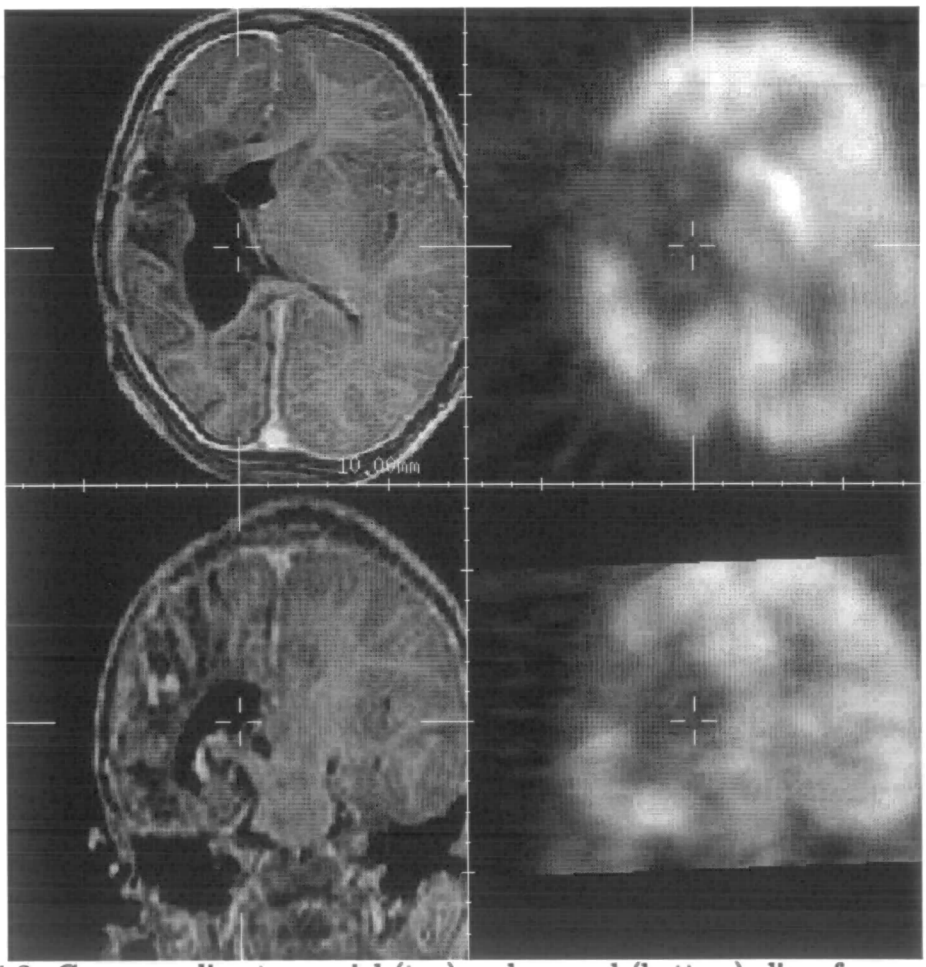


Figure 1.3: Corresponding transaxial (top) and coronal (bottom) slices from manually registered images showing different structures delineated by T1 weighted MR (left) and PET ^{18}FDG (right) scans of the brain.

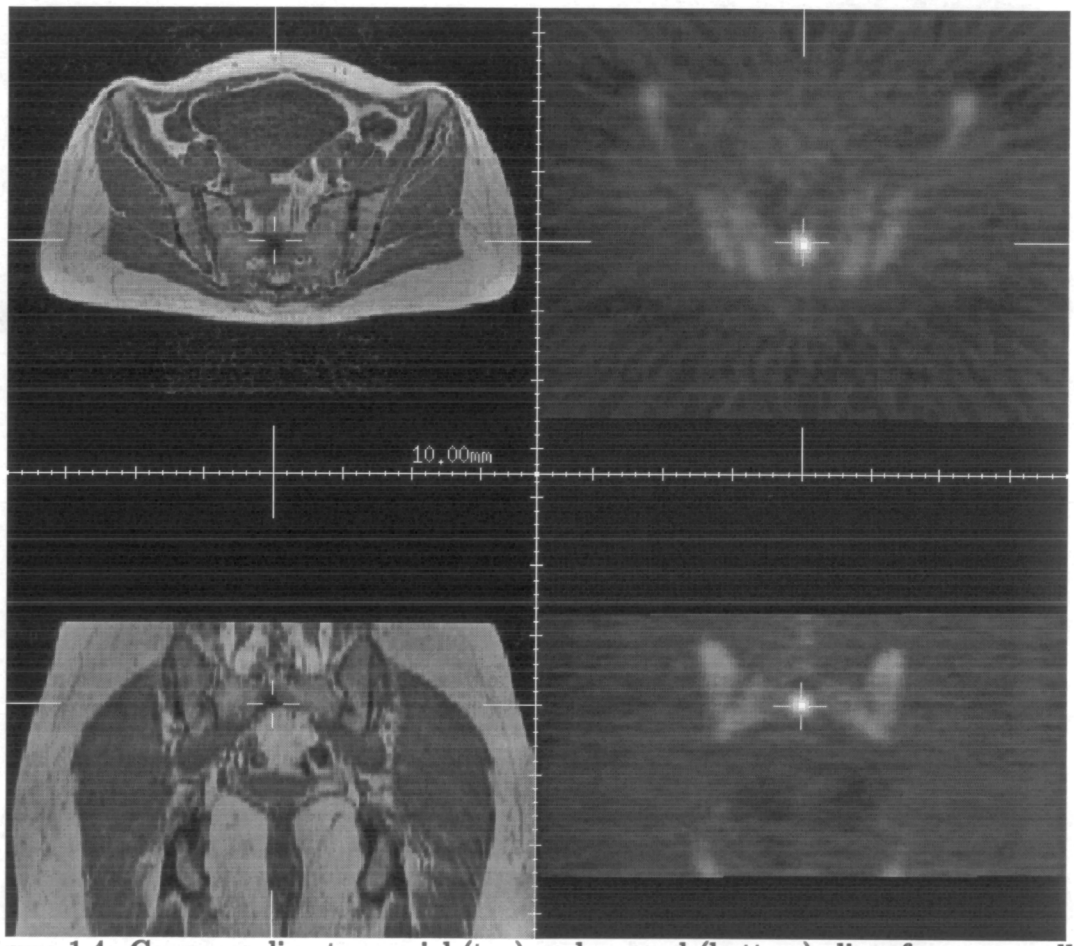


Figure 1.4: Corresponding transaxial (top) and coronal (bottom) slices from manually registered images showing different structures delineated by T1 weighted MR (left) and PET (right) scans of the pelvis. The PET image is of both ^{18}FDG and $^{18}F^{-}$ tracers which delineate both tumour and bone structure within the patient.

1.3 Hypothesis

One of the main aims of aligning different types of medical image is to enable us to accurately relate structures delineated by them. It is often the case that the more different the images, the more useful it is to be able to compare them when registered. Conversely, in order to constrain and define alignment, there must be shared structures in the two images. The approach taken in this thesis is that of deriving a *global* measure of image similarity which can be used as an indication of image alignment. By using all the data we hope to make best use of any shared structures within the images to provide a robust, accurate and therefore clinically useful image registration system.

The specific aims of the work in this thesis are as follows:

- Investigate and develop a 'voxel similarity' based description of multi-modality image alignment.
- Investigate current measures of alignment and develop new measures.
- Develop and implement a generic 3D image registration algorithm which optimises measures derived from an estimate of the joint probability distribution of image values.
- Experimentally compare image registration measures on a range of multi-modality image data to provide a generic image registration facility.
- Validate automated registration on a range of clinically typical image data to identify its limitations in clinical use.
- Extend entropy based measures of alignment using additional sources of information for cases where conventional measures fail.

1.4 An Overview of the Thesis

The thesis begins with a review of conventional approaches to image registration based on the extraction and alignment of corresponding features (chapter 2). This identifies the fundamental limitations of feature based alignment for fully automated registration.

The remainder of the thesis then divides into three main parts. The first part deals with the development of ideas behind measures of image similarity and alignment. The second part deals with their implementation and use for the task of brain image alignment. Finally, the last part presents extensions using information theory to deal with images for which conventional image similarity cannot provide a measure of alignment.

Image Similarity and Alignment

Chapter 3 develops a more generic description of alignment based on the volume of overlap of regions of material delineated by two modalities. From this a statistical description of alignment can be constructed based on the joint probability distribution or histogram of image values. The chapter then reviews in detail the development of similarity measures for multi-modality registration, from early approaches using pre-processing of data prior to correlation, to the development of more generic statistical measures of alignment.

Chapter 4 extends the statistical approach to alignment by developing the use of entropy as a measure of image alignment using techniques developed in information theory. Here we are treating the two modalities as sources of information about a scene and using information theory to quantify the shared structures. One of the most important aspects of these measures is their normalisation with respect to the local image statistics within the overlap of the two images. A simple model of a scene is used to investigate and compare the behaviour of different measures as the field of view of the image is varied.

Automated Brain Image Alignment

Chapter 5 then examines the practical aspects of evaluating and optimising image similarity measures between common 3D medical modalities. It breaks the task down into three main problems. The evaluation of corresponding measurements, the estimation of the joint histogram of image values and the optimisation of an alignment measure derived from this. A simple multi-resolution optimisation scheme is described which is used to compare the behaviour of similarity measures in the following chapters.

Chapter 6 is a detailed experimental comparison of the behaviour of the three main entropy derived measures from chapter 4, along with three statistical measures described in chapter 3. The investigation is based on a set of image data for which an accurate fiducial marker based alignment estimate is known. The experiments investigate the parameter space provided by different measures using random starts of the registration algorithm. The measures are compared with respect to varying image content, field of view and starting estimate of the registration. The results of varying transaxial field of view confirm the findings of the simulation experiments of chapter 4 and highlight the problems of normalisation of registration measures.

Chapter 7 and 8 look at the application of normalised mutual information to the fully automated registration of a large range of clinical image data. The aim here is to provide a fully automated registration system to replace current manual point based

image registration techniques in clinical use. The clinical image databases consist of 26 MR-CT image pairs of the head and 14 MR-PET image pairs of the brain. The automated estimates are compared to conventional manual registration estimates which provide an indication of clinically acceptable accuracy on data for which a gold standard is unavailable.

Introducing Additional Alignment Information

Chapter 9 develops the entropy based approach to image registration to enable the alignment of images containing severe intensity distortion. This is achieved using information theory to introduce extra spatial information about the intensities within distorted images, through a simple partitioning of the imaged space. The approach is applied to the registration of MR with PET images of the pelvis and surface coil MR angiography images of the head to CT and conventional head coil MR images.

Chapter 10 examines the limitations of voxel similarity based approaches. It describes an extension of the approach to recover the alignment of views where the most probable values in the two modalities do not correspond. This is again achieved using information theory to introduce additional information into the alignment process. In this case the additional information is derived from a connected component labelling of an image to distinguish between unconnected regions exhibiting the same values in an image.

Chapter 2

Identification and Alignment of Corresponding Features

2.1 Introduction

Medical image registration has been an active area of research for many years. The specific problem of multi-modality image registration has been an important aspect of this research for most of this time. There have been a number of recent reviews of image registration published. Brown [15] classifies types of registration transformation and approaches to estimating them, concentrating on the general registration problem. Van den Elsen *et al.* [96], Maurer *et al.* [68] and Lavallee [57] concentrated more specifically on the medical image registration problem.

In this thesis, it is intended to concentrate on the problem of *multi-modal* registration, specifically addressing approaches to the problem of deriving correspondence between different types of medical images. Approaches fall between two extremes. At one end, one can aim to specifically identify corresponding features in the two images, from which a direct measure of misalignment can be derived. Conversely a measure can be derived directly from the collective similarity between intensities (measurements) in the images, without assuming any direct correspondence.

This chapter reviews approaches based on the extraction and alignment of corresponding features. Many clinically usable solutions to these tasks are now in use, allowing automatic, semi-automatic and even fully automated alignment of various modality combinations for a number of applications. As a result, the problem is now often seen as being essentially 'solved' by a number of researchers. The aim of this chapter is to both summarise the body of work carried out prior to and during the study described in this thesis by other workers, and to justify the programme of work undertaken for this thesis.

2.2 The Spatial Relationship between Acquisitions

In order to describe the spatial relationship between two 3D image volumes we need a transformation T which relates points $\mathbf{y} = \{x, y, z\}$ in one image space say V_N to points $\mathbf{x} = \{x, y, z\}$ in the other, say V_M, so that,

$$\mathbf{x} = T(\mathbf{y}). \tag{2.1}$$

In practice the problem of aligning a region of a patient imaged in different scanners can be described by applying translations in three orthogonal axes and rotations around these axes. Additionally we may need to describe the relative scaling in each axes between the two scans, and the skew in each axis. This subset of transformations is specified mathematically using the constraints [15] that

$$T(\mathbf{x}_1 + \mathbf{x}_2) = T(\mathbf{x}_1) + T(\mathbf{x}_2),$$

and secondly that for all values of constant k,

$$T(k\mathbf{x}) = kT(\mathbf{x}).$$

These define the set of linear transformations. A further constraint, that the transformation

$$T_o = T(\mathbf{x}) - T(\mathbf{0}),$$

is linear defines the set of *affine* transformations. For all such transformations any straight line in one volume is transformed to a straight line in the other volume.

This includes only the Cartesian operations of translation and rotation, making up the the rigid transformation, and the scaling and skew factors. Such a transformation fully describes the relationship between images in the two modalities if all points in the patient remain fixed in relation to each other (i.e. the patient does not deform) between scans, and the patient is stationary throughout the acquisition period of the two images. In real terms the rigid transformation describes the difference in orientation of the patient in the two scanners, and the scaling and skew components describe the spatial acquisition parameters of the two scans.

The majority of work in multi-modality registration assumes that the scaling and skew parameters can be derived accurately from knowledge of scanner calibration. The multi-modality registration problem is generally to find the global rigid transformation between the two imaged spaces.

2.3 Corresponding Point Alignment

The most direct approach to registration is to first identify corresponding pairs of landmarks in the two images, and then to bring these points into alignment. This approach was first applied to the alignment of satellite images and can readily be applied to the alignment of anatomical landmarks in three dimensions. If we can identify some corresponding pairs of points within the two images, then we can consider each of these point pairs to be a solution to equation (2.1). If we have enough of these point pairs we can use them to find the value of T describing the alignment of the two images. Effectively, knowing the location of a subset of all pairs of points in the images allows us to find the transformation for the whole image. Mathematically the process of aligning the two sets of K selected points $X = \{\mathbf{x_i}\}$ and $Y = \{\mathbf{y_i}\}$, where $i = 1, \ldots, K$, can be expressed as simply minimising the mean square distance,

$$\mathcal{D}_p(T) = \sum_{i=1}^{K} \|\mathbf{x}_i - T(\mathbf{y}_i)\|^2, \tag{2.2}$$

between each pair. The task of finding a matrix T for this expression is known mathematically as the *Orthogonal Procrustes* problem, for which a number of approaches to a closed form solution have been described. One of the more popular approaches is that using a singular value matrix decomposition [3].

2.3.1 Landmark Identification

One of the most flexible and clinically applicable approaches to image registration has been that based on manual, interactive, point landmark identification [77]. The earliest work on registration [10, 31, 82] was initially carried out by the identification of points from printed films of the 2D slices. The availability of more powerful computing resources allows the use of interactive display tools to considerably simplify the process of point location [47]. Employing a flexible user interface [80] displaying orthogonal slices for point location enables clinically useful alignment to be achieved within 40 minutes by a trained user.

An obvious approach to automated registration is to develop a system capable of identifying known anatomical landmarks within images to remove the need for user interaction. Theoretical approaches to landmark identification have been developed [70]. In particular, Banerjee *et al* [6] employed geometric invariance properties. They extracted the entrance and exit points of concavities as invariant landmarks from 2D slices of MR and CT. In more recent work Rohr *et al.* [78] have investigated the use of differential operators to identify anatomical landmarks for non-rigid registration.

In general though, replicating the human action of recognising a specific anatomical structure may require a complex model based segmentation to ensure robustness to anatomical variation and image field of view. The task is made more difficult by the non isotropic sampling of many 3D medical modalities. This can mean that a small point structure is lost or distorted when imaged by thick slices in, for example, CT images. As a result, fully automated approaches using these techniques have found limited success.

2.4 Corresponding Surface Alignment

2.4.1 A Distance Between Corresponding Surfaces

An alternative, and more feasible task than the automated identification of anatomical landmarks, is the extraction of points on a corresponding boundary in both images. If it is possible to extract the same surface from the two modalities, then it is possible to derive an estimate of the alignment between the pair of images, as illustrated in Figure 2.1.

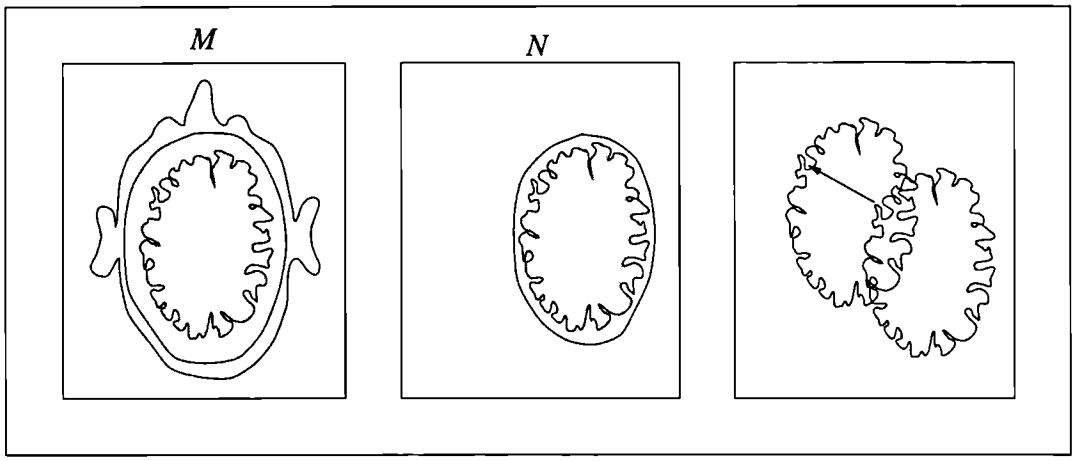


Figure 2.1: Extraction of the corresponding surfaces between which a measure of misalignment may be derived.

Surface matching can be thought of as the problem of aligning two sets of points where the one-to-one correspondence or pairing, is unknown. Given a pair of point sets,

$$X = \{\mathbf{x}_i\} \quad \text{where } i = 1, 2, \ldots K$$

$$Y = \{\mathbf{y}_j\} \quad \text{where } j = 1, 2, \ldots L$$

describing surfaces $\mathcal{X}$ and $\mathcal{Y}$ respectively, as illustrated by the outlines in Figure 2.2. Even if the surfaces exactly correspond, at registration the points on the surfaces are not necessarily aligned and, even the number of points on the two surfaces need not

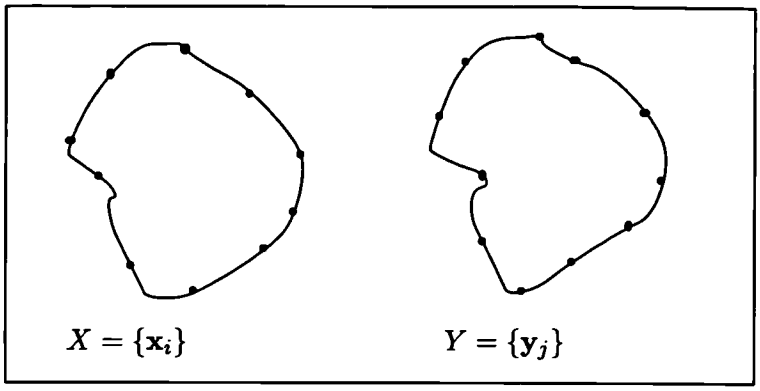


Figure 2.2: Outlines representing corresponding objects extracted from two modalities.

be the same. Because we do not have any point correspondence information, we need, for a given point $\mathbf{x}_i$, a function for the transformation T which returns a corresponding coordinate on the other surface described by the set Y, say $\mathcal{P}(T(Y), \mathbf{x}_i)$. This may simply be the nearest point in the second set or a more complex description (for example a point on the triangulation of surface points Y). Given this function, we can then develop a measure similar to that for point matching, based on the overall distance between the surfaces,

$$\mathcal{D}_s(T) = \sqrt{\sum_{i=1}^{K} \|\mathbf{x}_i - \mathcal{P}(T(Y), \mathbf{x}_i)\|^2}. \tag{2.3}$$

The presence of the point correspondence function $\mathcal{P}(T(Y), \mathbf{x}_i)$ means that, unlike equation (2.2), there can be no direct solution to equation (2.3), and so alignment by finding the value of T which minimises $\mathcal{D}_s(T)$ requires an iterative search.

Another way of looking at the difference between $\mathcal{D}_p(T)$ and $\mathcal{D}_s(T)$ is that $\mathcal{D}_p(T)$ is truly a measure of the distance from correct alignment of the two corresponding point sets. $\mathcal{D}_s(T)$ on the other hand, is not a direct 'distance to registration' but, simply a *measure* of the alignment of the two surfaces.

Where the descriptions of features are essentially determined by manual selection, the problem is one of developing an algorithm capable of finding the best alignment of the descriptions. The algorithm must be capable of aligning features which contain noise and outliers introduced by the observer or segmentation algorithm. The measure of registration is generally common to all approaches, the key tasks are:

- How do we identify corresponding surfaces ? (surface extraction)
- How do we improve a registration estimate? (optimisation)

- Which portions of the surfaces do we use to evaluate the distance ? (The outlier problem)

Approaches that have been proposed to addressing these problems are discussed in the following section.

2.4.2 Early Approaches

Pelizzari and co-workers [73] were the first to use surface matching to align different modalities. For MR and PET brain image alignment they employed semi-automated delineation of the external skin surface in each slice of the MR and PET scans to form a simple 3D surface description. For MR images, this consisted of the use of a simple threshold boundary following algorithm. The output of this was then manually edited. For the PET images, the corresponding surface was extracted in a similar way. Rather than use the clinical FDG image they traced the skin boundary from an edge enhanced version of the transmission scan. This is acquired clinically to attenuation correct the PET tracer image.

The skin surfaces were brought into alignment using a least squares fit, described by the authors as 'head-hat' matching. The correspondence function $\mathcal{P}$ can be computationally complex to evaluate, requiring an expensive search process. In order to make the task easier, the approach by Pelizzari employed a number of simplifications. For each point $\mathbf{x}_i$, the nearest point $\mathcal{P}(T(Y), \mathbf{x}_i)$ was taken as the point of intersection of a ray from the centroid of the hat points through $\mathbf{x}_i$, with the surface defined by Y. This approach assumes, given the centroids are initially aligned, that the two surfaces are predominantly spherical, so that the ray directions defined from the centroid are close to the true nearest point direction. This is a reasonable assumption for skin surfaces covering most of the head. It is less acceptable for more truncated volumes, or more complex surfaces such as the brain.

2.4.3 Surface Extraction and Correspondence

One of the inherent problems in applying surface matching to multi-modal image alignment is that of extracting a pair of corresponding surfaces from images of different physical measurements. Different workers have used a range of approaches to solving this problem. This section compares in detail the different approaches employed to extract and align corresponding features.

The task of corresponding object extraction is simplified in the early approach of Pelizzari [73], by the choice of the skin surface which is relatively easy to extract. The two problems posed by this are the accuracy of rigid alignment estimates due to skin

surface deformation, and the robustness of the matching due to the symmetries and lack of fine structure in the skin surface. The approach is dependent on the acquisition of a transmission scan as part of the attenuation correction of the tracer image. This imaging protocol is not always used in attenuation correction of PET imagery. In some cases, including our site, attenuation correction values are evaluated by the delineation of brain surface from the tracer image alone. There is a second problem with using the transmission scan, which is the assumption that the transmission and emission scan are themselves in perfect alignment. In many cases this is not guaranteed, and the amount of movement that can occur is dependent particularly on the time between transmission and emission acquisitions in the clinical protocol. These issues will be discussed in more detail in chapter 9 in the alignment of PET images of the pelvis with other modalities.

For MR and PET registration Oghabian and Todd-Pokropek [72] employed the same approach as Pelizzari, using skin surfaces extracted from the MR and the transmission scan of the PET image. For registration with HMPAO SPECT imagery, where skin surface extraction was more difficult, the brain surface was used. In order to make the extraction of brain surface from MR easier, a fat-suppressed STIR sequence was employed to reduce the strength of the skin surface signal. Some user selectivity in the choice of regions of surface used in registration was employed to avoid the major differences in the modalities. An example is their exclusion the eyes from an MR head image where the SPECT data does not adequately represent the anatomical complexity.

Turkington *et al.* [91] employed manually initiated extraction of the brain boundary from MR and PET (FDG and H_2O). A polygon was manually drawn around the brain in each slice and the brain edge within this region defined by a simple threshold. The MR boundary was extracted at MR resolution and then smoothed to simulate the lower resolution contour available from PET. Smoothing was a simple 2D averaging of the radial distance to the centroid of each contour within an axial slice. Transaxial sampling was reduced by skipping alternate slices in the contouring process. A similar approach was used for PET contour extraction but in order to avoid problems caused by variable tracer uptake in PET, an algorithm to identify the first derivative maximum was employed using a smoothed derivative kernel. The authors report manual editing was sometimes required. The same approach was used for H_2O PET studies but because of the higher noise level, the images were first smoothed. The authors indicated that the brain surface extracted from H_2O PET would contain vascular structures that were not present in either the MR or PET FDG brain surfaces and this partially explained the poorer results for MR-H_2O alignment.

The work of van Herk [97] is important because it specifically proposes a fully auto-

mated registration scheme using automated corresponding surface extraction and matching. The work contains both a detailed description of the method and an evaluation on a significant number of clinical datasets. A detailed examination of the work highlights a number of limitations in the surface matching approach when applied to multi-modal image registration.

The method employs a number of heuristics and assumptions about the images to enable the automated extraction of corresponding surfaces from the different modalities. Object detection in CT is achieved firstly by setting a simple bone threshold, the skull boundaries are then selected from other spurious boundaries by setting a minimum threshold of boundary length below which objects are ignored. Separation of points on the inner surface of the skull from points on the outer surface is achieved for each slice individually by finding the location of the points with respect to their centre of gravity, grouping into angular ranges around this and picking the nearest point to the centre. This effectively assumes a spherical distribution of points and limits the approach to predominantly axial CT scans covering a large proportion of the skull.

In addition to this, the approach assumes a simple relationship between the inner skull boundary from CT and the brain boundary extracted from SPECT. To account for the discrepancy in correspondence, the CT boundary was first dilated by 1mm. This assumption may be a significant simplification of the relationship between skull and brain boundaries. In practice a number of practical limitations will apply including:

- Uniform distance will not apply over the entire brain surface, for example there will be a difference between the distance at points on sulci and gyri.
- The distance may vary between patients.
- The distance may vary with the medical condition of the individual.

For MR-CT registration, skull boundaries were extracted from MRI for matching with those from CT. A rough representation of these were extracted by histogram thresholding at 15% and 65% of the maximum MR value. The authors make the point that this produced a low quality segmentation, which was adequate to achieve registration. In practice the tests were limited to T1 weighted and proton density MRI images from one scanner. Unlike CT, the relationship between absolute MR values and tissue properties can vary considerably between different MR sequences and even between the same sequence on different MR machines.

In their MR-CT surface matching algorithm, Maurer *et al.* [67] were aware of the surface correspondence problem. In their approach using T2-weighted MR imagery

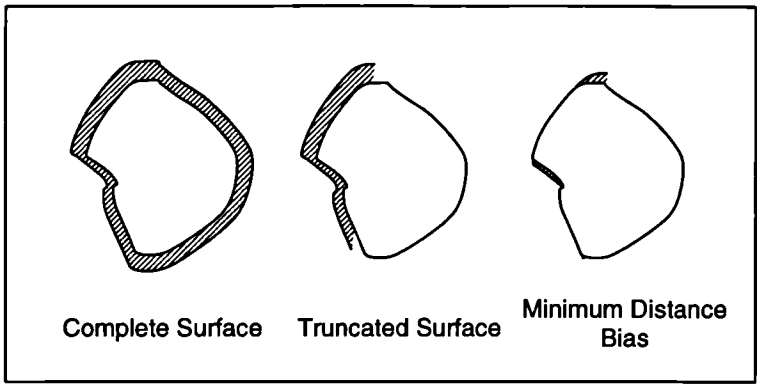


Figure 2.3: An illustration of the effect of minimising the distance between related but not directly corresponding boundaries, when the boundary is not complete.

they aligned the inner surface of the skull from CT with boundaries representing the cerebrospinal fluid (CSF)-dura interface from MR. The CSF boundary is much more closely related to the inner skull boundary, but there are a number of areas where the two are appreciably different particularly where there are blood vessels between the dura and skull. This is visible in their illustrated results showing a large visible discrepancy between the CSF boundary and the skull around the sagittal sinus.

Improved Approaches to Optimisation

One of the main computational problems faced in matching complex surfaces represented by many points, is the task of finding the nearest point $\mathcal{P}(T(Y), \mathbf{x}_i)$ on the other surface. This effectively means searching through all points in the set of points Y to find those nearest to $\mathbf{x}_i$. One of the most successful and elegant approaches to simplifying this problem was proposed by Barrow [7] and Borgefors [11]. This employs a *distance transform* in the evaluation of the distance between corresponding boundaries.

Rather than represent the two surfaces to be matched simply by two lists of points, one of the objects delineated can be stored as a binary image where voxel or pixel values are either object or background. From this binary image, a *distance image*, can be formed using a 'distance transform' where each voxel or pixel outside the object is given a value representing its distance to the nearest point on the object. To evaluate the overall distance between two objects for a given T, the set of points describing the object in one image are simply transformed to locations in the distance image of the other object. The voxel values then give a direct estimate of $\mathcal{P}(T(Y), \mathbf{x}_i)$ for each point. This precalculation of a discrete form of $\mathcal{P}(T(Y), \mathbf{x}_i)$ offers appreciable speed improvements.

An important difference in the matching process is the type of distance transform employed. It is possible to evaluate a full Euclidean distance transform, but this can be

computationally intensive. An approximation to this is the chamfer distance transform which can be produced using a set of kernels applied to the binary 'object' image. One of the first researchers to apply this technique to medical image registration was Jiang *et al.* [52].

An alternative approach to simplifying the computational task was taken by Besl and McKay [8]. Rather than reduce the computational complexity of evaluating $\mathcal{P}(T(Y), \mathbf{x}_i)$, they attempted to reduce the number of times $\mathcal{P}(T(Y), \mathbf{x}_i)$ needs to be evaluated. They did this by developing an improved optimisation scheme for $\mathcal{D}_s(T)$, the so called *iterative closest point* algorithm. This takes advantage of the closed form solution to the corresponding point matching problem. Each iteration can be split into two steps, first a correspondence function $\mathcal{P}(T(Y), \mathbf{x}_i)$ returns the closest coordinate for each $\mathbf{x}_i$ for the initial estimate T_1, as with a direct iterative search. An improved transformation T_2 is then evaluated directly using the closed form solution of equation (2.2). This provides a much improved estimate for T with only one evaluation of corresponding surface points.

The main drawback of the approach is the need to interpolate truly nearest points on the surfaces to achieve a useful solution from the point alignment step. Since the two point sets X and Y are not directly corresponding, $\mathcal{P}(T(Y), \mathbf{x}_i)$ must return a true closest coordinate on $\mathcal{Y}$ rather than simply the closest point in Y as illustrated in Figure 2.4. Producing an accurate estimate of this closest coordinate is appreciably more complex and this, to some extent, offsets the improvements provided by the reduction in the number of evaluations of $\mathcal{D}_s(T)$ required.

A similar approach to this was proposed by Chen and Medioni [17]. Rather than minimise the distance between closest points, they used the distance to the corresponding tangent plane of the surface. In a similar approach to Besl and McKay the distance was reduced by iteratively evaluating a least squares solution. This was used to successfully align multiple, overlapping range images.

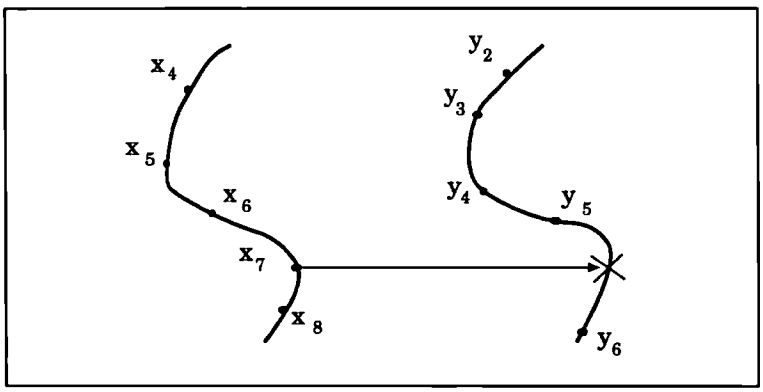


Figure 2.4: Function $\mathcal{P}(T(Y), \mathbf{x}_i)$ returning the true closest point on surface $\mathcal{Y}$

Oghabian and Todd-Pokropek [72] employed a multi-resolution evaluation and optimisation scheme to avoid local minima when aligning surfaces. In addition they employed an ordering and grouping of the surface points based on location and direction of their surface normals to simplify the alignment process [71].

2.4.4 Outliers and Surface Extent

Another important aspect of matching surfaces from different scanners is that of handling outliers produced by poor segmentation or differing field of view. One surface may cover a much larger portion of the body than the other. Alternatively the two scans may cover significantly different portions so that their region of true correspondence, say $\mathcal{R} = \mathcal{Y} \cap \mathcal{X}$, is only a small portion of both scans. Unfortunately we don't know the region of correspondence until we know the registration solution.

If the function $\mathcal{P}(T(Y), \mathbf{x}_i)$ allows contributions from every part of both surfaces for a given T, then the error measure may contain large contributions from points at the extremity of the two scans. At registration we are then including displacements between points that have no real correspondence, as illustrated in Figure 2.5. This can produce a range of effects on the behaviour of $\mathcal{D}_s(T)$. It may simply produce a more complex function of misregistration with more local minima, making the task of finding the correct registration difficult. In the extreme case it may actually produce a distinct global minima not corresponding to true alignment, and remove any form of minima at true registration.

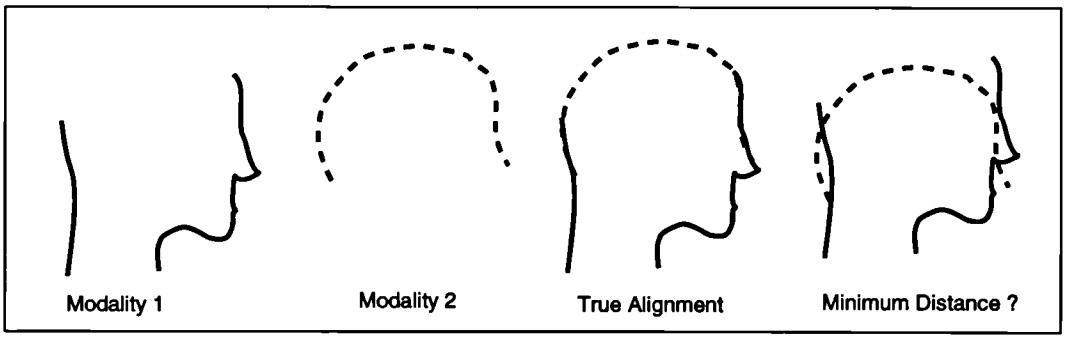


Figure 2.5: The effects small volume of overlap on minimisation of surface distance.

A related problem is that of outliers produced by local differences in the object boundaries delineated in the two modalities, resulting in portions of one surface that do not correspond to structures on the other. These could shift the correct optimum leading to errors in alignment, as illustrated in Figure 2.6. One approach to this problem is to specifically exclude regions where differences occur, but again this generally requires some form of manual intervention.

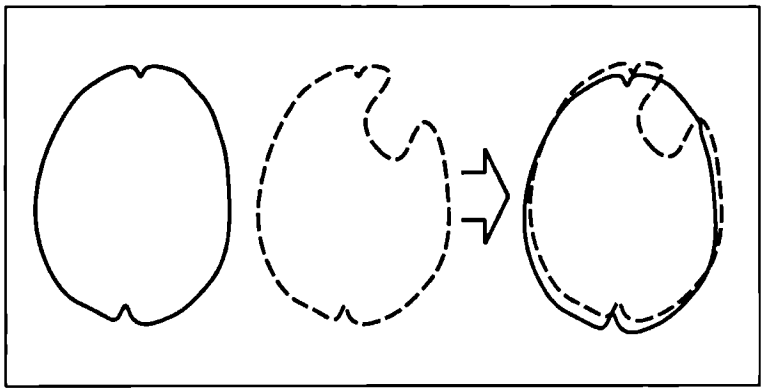


Figure 2.6: Outliers caused by localised differences in the delineated surface.

The handling of outliers has received less attention in much of the surface matching literature. It's importance depends very much on the application. Jiang *et al.* [52] employed a simple form of thresholding, to exclude the influence of outliers, in their adaption of the chamfer matching algorithm of Borgefors. Distances to nearest points were simply thresholded to a particular value chosen experimentally. This threshold was applied directly to the values in the distance image used in their chamfer matching implementation. They derived the level of threshold as a fraction of the "match surface dimension". For MR-PET brain image matching they quote a threshold level of 2% but no details are given of levels for other image types or anatomical regions.

The work of Geraud [40] introduced a more complex adaptive distance threshold derived from the distribution of surface distances. The approach appears to provide some improvements but at a high computational cost.

The most comprehensive investigation of these problems on a range of data has been published by Turkington *et al.* [91]. This work examined the use of the surface matching approach proposed by Pelizzari *et al.* [73] for MR and PET brain image registration when presented with truncated surfaces. Scans of normal volunteers were aligned using markers and brain surfaces extracted from the two modalities as described in section 2.4.3. Truncation of the PET surface was carried out by three approaches, firstly using only the left half of the brain, secondly the left half of the brain plus one full slice including the right of the brain and finally the central 6 slices (39mm) of the scan. They present the results for six (rigid) and nine (rigid and scaling) parameter matching of the surfaces, which show the quality of the *surface match* by the reproducibility of the scaling estimate. The comparison of the estimates produced using the full surface to those produced using partial surfaces are much less convincing with appreciable differences in rotational estimates (up to 4 degrees in some cases). No results or discussion on how robust the registration is to the quality of the starting estimate was included.

2.4.5 Using Related Surfaces

The approaches described so far assume that extracted surfaces are directly corresponding. In many cases it is not possible to accurately extract corresponding surfaces from two different modalities (a surface may be missing completely from one modality or simply have very low contrast). It may though be possible to accurately extract neighbouring surfaces between which there is some known spatial relationship. A classic example is that of the outer brain surface and inner skull surface extracted from MR and CT respectively. The distance between these surfaces can have two important limitations as a measure of alignment. Firstly it may poorly constrain the registration (there may be a range of solutions where this is equally minimum due for example, to symmetries in the surfaces). Secondly, if there is a distinct minimum it need not correspond to correct alignment of the two surfaces.

The work of Hill [48] was the first to address this problem enabling the use of neighbouring surfaces to derive a registration estimate. This was achieved by introducing information about the expected distribution of distances between points on the two surfaces. Registration was achieved by seeking the transformation which provided the correct distribution of these distances. This work illustrated registration using alignment of the inner skull surface extracted from CT and the outer brain surface extracted from MR.

One possible problem with this approach is that the expected distribution of distances is itself a function of the overlapping regions of the two surfaces. This distribution will thus vary between image pairs with different fields of view and cannot be determined prior to alignment.

2.5 Combining Different Types of Geometrical Features

One of the major criticisms of feature based approaches is that they are using only a small part of the information available in the images to determine image alignment. In many cases a surface may not constrain the registration adequately or there may be inadequate numbers of corresponding points in the images to provide an accurate estimate. Recently there has been considerable interest in using different types of geometrical features, such as points and lines in determining 3D alignment. These approaches can be divided into two classes. There are those using information from different features *sequentially* in the registration process. There are then those approaches employing a mathematical combination of distances between different features to align them all *simultaneously.*

2.5.1 Sequential Feature Alignment

Some work has been done on simplifying the manual registration task by dividing it up into 2D registration steps. The work of Kapouleas *et al.* [53] employs an interactive alignment of the inter-hemispheric fissure in MR and PET images of the brain. End points of the fissure are identified in each slice of the two scans to form descriptions of the two planes which are then aligned. Edges are automatically detected in the MR image highlighting boundaries of anatomical structures. This is achieved by the application of a 2D edge detection kernel followed by simple thresholding to create a simplified boundary image. This boundary image is then overlaid onto sagittal PET images to visualise the alignment of the two. Interactive modification of translation and rotation parameters in the sagittal planes then allows the full 3D alignment to be solved. It is interesting to compare this approach to that of surface and point matching. In those methods a manual segmentation or identification of features is used followed by an automated alignment of the structures. In this method essentially an automated identification of features is employed for visualisation followed by manual alignment.

The work of Ge *et al.* [39] is closely related to this approach. A similar alignment of the inter-hemispheric fissure is employed to achieve alignment of the sagittal plane. The user specifies corresponding anatomical landmarks from which the transformation parameters are estimated. All of these sequential methods are essentially intended to aid manual alignment.

2.5.2 Simultaneous Feature Alignment

The work of Collignon *et al.* [20] was arguably the first to employ a measure derived from both points and surfaces simultaneously. The final distance measure is a combination of point distances and surface distances calculated separately. In order to optimise the function they employed a generic multi-dimensional optimisation scheme derived from that described by Powell [74].

Meyer *et al.* [69] described an approach to combining alignment distances from points, lines and planes in one measure. This is achieved by reducing lines to a pair of points and planes to a single point. The influence of each feature is separately weighted by its variance and the final result is estimated directly using the single value decomposition approach to point matching. The approach appears to simplify and ignore valuable alignment information provided by lines and surfaces. No comparison with direct point matching or surface matching alone is provided in the results to indicate increased accuracy. The main advantage in the approach is in dealing with image pairs which have few corresponding points but where lines and planes may be identified.

Maurer *et al.* [67] presented the most general approach to feature alignment by extending the iterative closest point algorithm of Besl and McKay [8]. The approach uses a weighted measure of the distance between geometric features. They demonstrated the approach for the registration of MR and CT brain images using points and surfaces combined. They compared the results obtained to using points and surface alone and showed a small but statistically significant improvement in accuracy when compared to a gold standard derived from fiducial marker points.

2.6 Generic Feature Alignment

Corresponding surface and point matching rely on the identification of specific corresponding structures within the two images. Rather than attempt to extract corresponding occurrences of a particular geometrical feature (e.g. one surface), it easier to automatically extract all occurrences of a feature from both modalities. By choosing a particular class of geometrical feature present in both data sets we can simplify the alignment process.

Such an approach is used by Thirion *et al.* [89] who extract a class of 3D lines called *crest lines* from a volume image. These 3D lines are the equivalent of corner points of objects in a 2D image and can be extracted from a segmented volume using an efficient marching lines algorithm. The lines are brought into registration using a geometric hashing algorithm [44].

The work of van den Elsen *et al.* [93, 94] is closely related to this. Here, particular types of geometrical feature present in both MR and CT are selected. They then employ a voxel based approach to the alignment of images of the strength of detected features. This will be discussed in more detail in the following chapter.

2.7 Voxel Based Object Alignment

An alternative to extracting a description of object boundaries is to retain a binary image based description of an object. Rather than use a distance between object surfaces as the measure of alignment, we can form a measure directly from the match of the labelled volumes. There has been considerable work in the area of binary and grey scale object alignment using a number of related techniques. These approaches are generally assuming a simple image of an object where all features correspond.

Faber and Stokely [33] compared the use of tensor based moment and principle axes methods to recover object alignments. Essentially both of these methods rely on the decomposition of translation and rotation components and the initial solution of the

translation components by alignment of the centre of mass of the objects. Both methods employ the evaluation of the moments of an object. The principle axes are those for which the moments of inertia of a shape are minimised. By evaluating the direction of these axes for two objects, the relative rotations between the axes provides an estimate of their rotational misalignment.

The authors examined the behaviour of the approach to varying resolution of both binary and grey scale 3D images of a synthetic object and a volume blood pool SPECT image. The translation, rotation and scale parameters were estimated and the results analysed by evaluating the correlation and difference between the original and estimated images. Their results indicated superior performance by the principle axis method on lower resolution images. The results were essentially limited to within modality imaging.

One of the main limitations of these approaches is the effect of truncation of the shape of an object by a limited field of view in one or both images, introducing considerable errors in the initial estimation of the centre of mass. Dhawan *et al.* [27] tackled this problem by developing an iterative principle axes method to align PET images with a smaller axial extent, to manually segmented MR of the brain. The approach was designed to handle the alignment of the smaller volume PET images to the larger volume MR by iteratively computing the centroid and principle axes over a decreasing field of view.

These methods are closely related to those of surface matching and Rusinek *et al.* [81] compared the behaviour of surface matching and principle axes methods to align images of the brain derived from MR and CT.

2.8 Summary

This chapter has covered the large area of work based around the approach of registration by corresponding feature alignment. As a review it is by no means complete but provides an idea of the types of solution and their limitations. The approach is attractive for many reasons. One of the most important in the past has been computational efficiency (both RAM and CPU) from the use of compact representations of the image structure.

One of the key problems in this approach is the task of extracting corresponding features (surfaces, lines or points) from different medical modalities. In many of the approaches, this problem is either handled by significant manual intervention, or solved approximately by simple heuristics. Any generic automated solution must include knowledge of imager characteristics and the anatomy and physiology involved in the imaging process.

Inherently, the accuracy and applicability of the technique to clinical images is limited by the fact that registration is derived from a very small portion of the image

data. This becomes a significant problem when applying the approach retrospectively to the alignment of truncated image volumes, where there are limited numbers of easily extractable corresponding features. This limitation has been partially addressed by techniques to combine alignment measures from different types of geometrical features either sequentially or simultaneously in the registration process.

The problem of automated corresponding feature extraction limits both the accuracy and robustness of the technique in clinical practice. Approaches extracting all examples of a particular class of higher level geometrical feature are one way of avoiding the complex task of corresponding object detection. These approaches though have not yet been shown to be robust to differences between multi-modality image pairs.

Chapter 3

Image Similarity As a Measure of Alignment

3.1 Introduction

In the previous chapter approaches based on the extraction and alignment of corresponding features were reviewed. An alternative approach to minimising a distance between corresponding features is to maximise a measure of *similarity* between two images. We can express an overall image similarity $\mathcal{S}$ for a transformation T between the two images, as sum of individual measurement similarities $S(.)$, at Q different points,

$$\mathcal{S}(T) = \sum_{i=1}^{Q} S(m(\mathbf{x}_i), n(T(\mathbf{x}_i))), \tag{3.1}$$

where $m(\mathbf{x}_i)$ and $n(T(\mathbf{x}_i))$ are corresponding values. This approach can be related back to that of feature alignment equation (2.3). We would like $\mathcal{S}$ to be a monotonic function of misalignment, which behaves like a mean distance between features. For this to be the case, the number of similar pairs ($\sum_{i=1}^{Q} S(.)$) should fall smoothly with misalignment. We must therefore assume that there are uniform regions of corresponding values in the two images at alignment. In many cases this is a valid assumption for many 3D medical modalities. The key question in this approach is then how do we say whether values in the modalities are *similar*, i.e. what do we use for our function $S(.)$?

This chapter starts with a description of multi-modality imaging and alignment in terms of overlapping regions of material delineated in the two modalities. From this a statistical view of the alignment is developed which summarises the relationship between these delineated regions. This forms the basis for a review of the development of simple statistical measures of image alignment.

3.2 Views of Anatomy and Physiology

Multi-modality imaging is the process of using measurements of different physical properties to delineate different regions of material in the body. We can refer to a single modality as providing a *view* of the underlying *scene*. It is the unknown relationship between these views which forms the basic problem of multi-modality alignment.

In this case, unlike conventional computer vision problems, the view of the scene differs not because of camera angle or the removal or introduction of objects, but simply because of the difference in the property being imaged. A related field is that of multi-spectral remote sensing, where it is conventional for images at different wavelengths to be acquired simultaneously and therefore in correspondence. In medical imaging such an arrangement is unlikely to be feasible either technically or financially in a clinical setting.

3.2.1 Materials, Properties and the Delineation of Regions

To look at the problem we can start by taking a simple description of the scene to be imaged. This consists of a volume of material say $V_M \subset \mathbb{R}^3$, for which we have some set of anatomical or physiological classifications $\tilde{R}$. These classifications partition up the volume into a set of regions $r \in \mathcal{R}$, each made up of points $\mathbf{x} \in r$ in space. More formally, each of these regions for our purposes internally exhibits uniform values with respect to measurements of any property. The regions cover the whole space,

$$\bigcup_{r \in \mathcal{R}} r = V_M, \tag{3.2}$$

and are non-overlapping (disjoint) so

$$\forall s, r \in \mathcal{R}, s \neq r \leftrightarrow r \cap s = \emptyset. \tag{3.3}$$

Note that each point in a region $r \in \mathcal{R}$ need not be spatially connected to the rest of the points, r is just the set of points with the same classification.

We can image the scene with a device which will make measurements of a property at points $\mathbf{x} \in V_M$. The device will record the actual underlying values $\tilde{m}(\mathbf{x})$ as an image of measurements $m(\mathbf{x})$. Here a value falls within the set $\tilde{M}$ of possible values and a measurement within the set M of possible measurements. Examples of properties imaged for medical use are:

- X-ray linear attenuation coefficient μ
- Radioisotope density resulting from tracer uptake in tissues
- Proton spin density and relaxation times

In the following discussion there is an important distinction between the actual value of a property exhibited by a material, and the *measurement* of that value. The multi-modality alignment problem can be described in terms of the differences in the underlying values in a scene, although differences in measurements of those values due to the way images are formed will be covered in later chapters.

A scene may be divided into sets of points (which we will term regions) which exhibit a particular value,

$$\psi(\tilde{m}) = \{\mathbf{x} | \mathbf{x} \in V_M, \tilde{m}(\mathbf{x}) = \tilde{m}\}. \tag{3.4}$$

The complete image will consist of the set of regions delineated by each of the occurring values $\tilde{m} \in \tilde{M}$,

$$\mathcal{M}(\tilde{M}) = \{k | k = \psi(\tilde{m}), \tilde{m} \in \tilde{M}\}. \tag{3.5}$$

The regions $\mathcal{M}$ are determined by the properties $\tilde{m} \in \tilde{M}$, of each of the anatomical or physiological classifications $\tilde{r} \in \tilde{R}$ summarised by the mapping,

$$\phi : \tilde{R} \mapsto \tilde{M}. \tag{3.6}$$

If there exists a simple injective (one-to-one) mapping from classification of material to properties of those materials, then we have an ideal modality which can delineate all the regions of interest within the patient. In practice this is often not the case, and the imaged property does not distinguish all the regions needed for a clinically sufficient description. Here there will be one or more pairs of regions $r \in \mathcal{R}$ which exhibit the same (or at least, indistinguishably equivalent) values of the property so that the mapping ϕ is not injective.

As a result a second modality is often used to provide an image $\tilde{n}(\mathbf{x})$ of measurements of a second property $\tilde{n} \in \tilde{N}$, for which the mapping $\phi : \tilde{R} \mapsto \tilde{N}$ will be different. This new image will then delineate a second set of regions, $\mathcal{N}(\tilde{N})$ as illustrated in Figure 3.1.

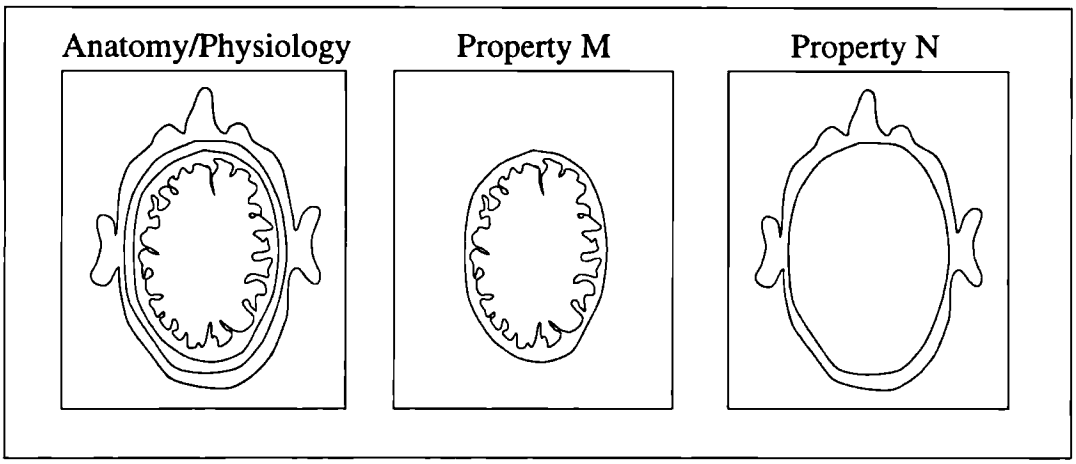


Figure 3.1: Alternative 'views' of structure within a volume provided by measurements of different properties.

3.2.2 Combining Views of a Scene

In general we can produce a single merged set of regions $\mathcal{F}$ from the two modalities by intersecting each of the regions delineated by the two images,

$$\mathcal{F}(\mathcal{M},\mathcal{N}) = \{k \cap l | k \in \mathcal{M}, l \in \mathcal{N}\}. \tag{3.7}$$

These regions may alternatively be arranged as a 2 dimensional array of sets,

$$\mathcal{F}(\mathcal{M},\mathcal{N}) = \begin{bmatrix} k_1 \cap l_1 & . & . \\ . & . & . \\ . & . & k_i \cap l_j \end{bmatrix} \tag{3.8}$$

where i and j are the number of distinct values of the properties M and N respectively.

There are three ways in which the two properties can delineate a region. Firstly, a region may be delineated by both modalities (region r_3 in Figure 3.2). This results in a single none empty set in a row and a column in the 2D array,

$$\mathcal{F}(\mathcal{M},\mathcal{N}) = \begin{bmatrix} . & \emptyset & . \\ \emptyset & k_2 \cap l_2 & \emptyset \\ . & \emptyset & . \\ . & \emptyset & . \end{bmatrix}. \tag{3.9}$$

Alternatively the region may be delineated by only one of the properties (region r_2 in Figure 3.2), resulting in a single none empty set in a row or column,

$$\mathcal{F}(\mathcal{M},\mathcal{N}) = \begin{bmatrix} k_1 \cap l_1 & \emptyset & . \\ \emptyset & . & . \\ k_3 \cap l_1 & . & . \\ \emptyset & . & . \end{bmatrix}. \tag{3.10}$$

Finally the region may only be delineated as an intersection of regions in the two modalities, resulting in multiple none empty sets in both a row and a column.

$$\mathcal{F}(\mathcal{M},\mathcal{N}) = \begin{bmatrix} . & . & k_1 \cap l_3 \\ . & . & \emptyset \\ . & . & \emptyset \\ k_4 \cap l_1 & \emptyset & k_4 \cap l_3 \end{bmatrix} \tag{3.11}$$

This is illustrated by region r_5 in Figure 3.2.

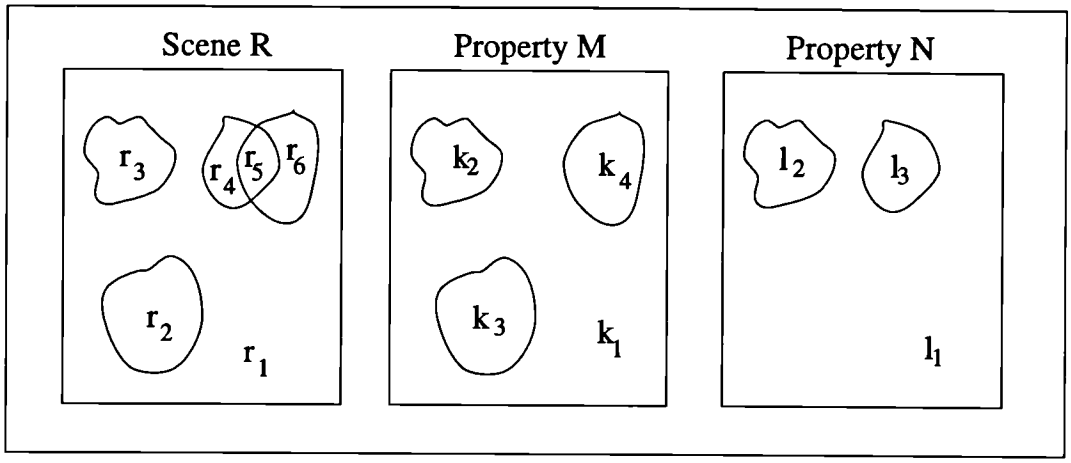


Figure 3.2: The delineation of regions in a scene by different physical properties.

3.2.3 Spatial Correspondence

In practice the process of acquiring multiple images of different properties cannot generally be done simultaneously or, even worse, using the same imaging system (or hospital). Mathematically the second set of delineated regions $\mathcal{N}$ is a partitioning of a different space $\mathcal{N} \subseteq \wp(V_N)$, and the mapping or transformation,

$$T : V_N \mapsto V_M \tag{3.12}$$

is unknown. For any given transformation T we get a set of delineated regions $\mathcal{F}(\mathcal{M}, T(\mathcal{N}))$. The problem of registration is to find the transformation T_R that gives us the correct description of the anatomical or physiological regions in the patient. To do this we need a way of defining when

$$T_R = T : \mathcal{F}(\mathcal{M}, T(\mathcal{N})), \tag{3.13}$$

in other words we need a *registration criteria.*

3.2.4 Registration Criteria

One plausible indication of alignment of regions $k \in \mathcal{M}$ and $l \in \mathcal{N}$ in the two modalities is the volume of their intersection, say $\nu(k \cap l)$. We can consider the matrix of volumes

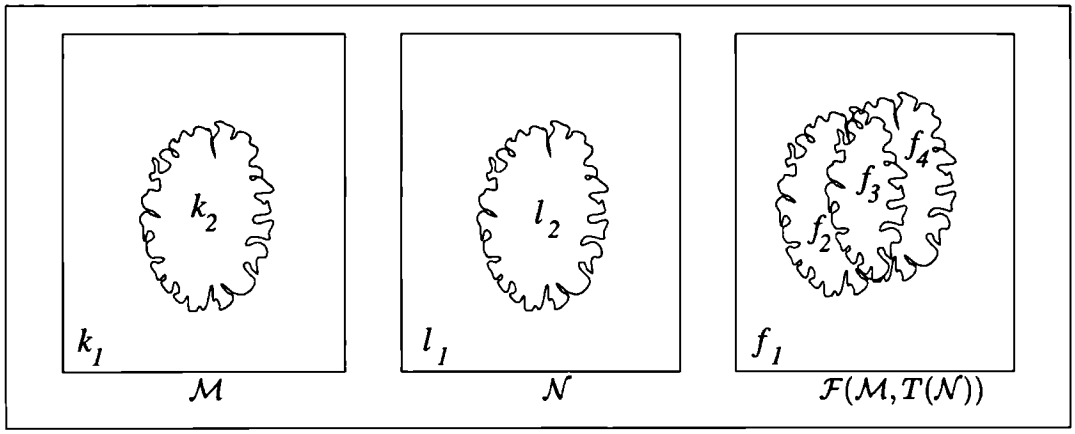


Figure 3.3: A spatial transformation T between two modalities delineating regions within the patient provide a combined set of regions $f \in \mathcal{F}$.

of all the intersecting regions say,

$$U(T) = \begin{bmatrix} \nu(k_1 \cap T(l_1)) & \nu(k_1 \cap T(l_2)) & \dots & \nu(k_1 \cap T(l_j)) \\ \nu(k_2 \cap T(l_1)) & \nu(k_2 \cap T(l_2)) & \dots & \nu(k_2 \cap T(l_j)) \\ \cdot & \cdot & \cdots & \cdot \\ \nu(k_i \cap T(l_1)) & \nu(k_i \cap T(l_2)) & \dots & \nu(k_i \cap T(l_j)) \end{bmatrix}. \tag{3.14}$$

Each volume will be a function of the transformation T. As $T \mapsto T_R$ there will be some regions, say k_α and l_β, in $f \in \mathcal{F}(\mathcal{M}, T(\mathcal{N}))$ which disappear i.e.,

$$\lim_{T \mapsto T_R} \nu(k_\alpha \cap T(l_\beta)) \mapsto 0. \tag{3.15}$$

Alternatively, simply the rate of change of their volume of intersection may tend to zero,

$$\lim_{T \mapsto T_R} \frac{d\nu(k_\alpha \cap T(l_\beta))}{dT} \mapsto 0. \tag{3.16}$$

Because T is effectively a vector of 6 parameters (for rigid alignment), there may be some directions for which a transformation results in a change in volume and others for which it does not. We can express this mathematically using directional derivatives as,

$$(\delta \mathbf{T}.\nabla)\nu(k_\alpha \cap T(l_\beta)) = 0, \tag{3.17}$$

where

$$\nabla = \left(\frac{\partial}{\partial t_1}, \dots \frac{\partial}{\partial t_6}\right), \tag{3.18}$$

and the direction of the vector

$$\delta \mathbf{T} = (\delta t_1, \delta t_2, \delta t_3, \dots) \tag{3.19}$$

determines the direction of misalignment.

Where the volume tends to zero for all directions there are three cases of interest. A region may simply reach a maximum, or minimum volume at registration. These are illustrated for a pair of very dissimilar images approaching a possible 'alignment' in Figure 3.4. The third case describes a rather more subtle alignment where there is

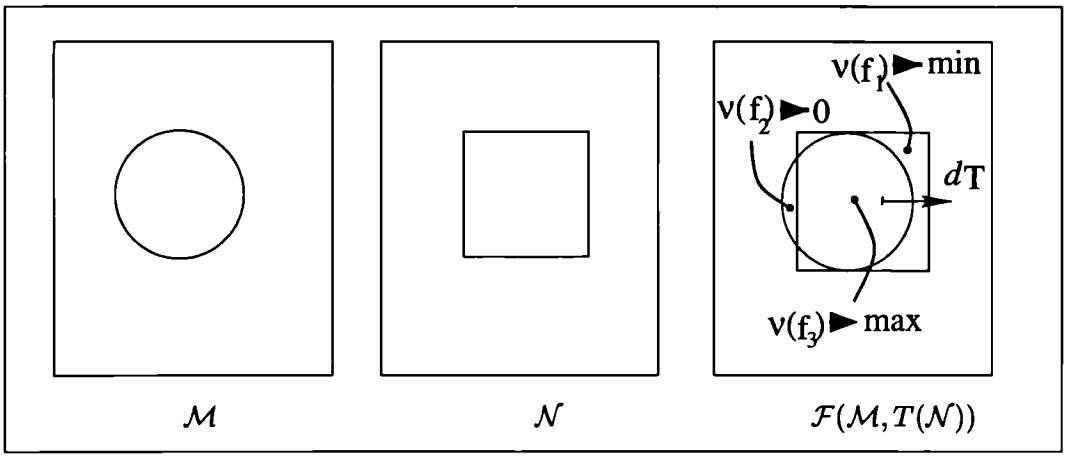


Figure 3.4: An illustration of the changing volume (area) of intersecting regions (right), delineated in different modalities (left and centre), as the transformation between the two modalities approaches 'alignment'.

a point, corresponding to the alignment of part of the border of a region with another as illustrated in Figure 3.5. This case in itself is not enough to constrain alignment

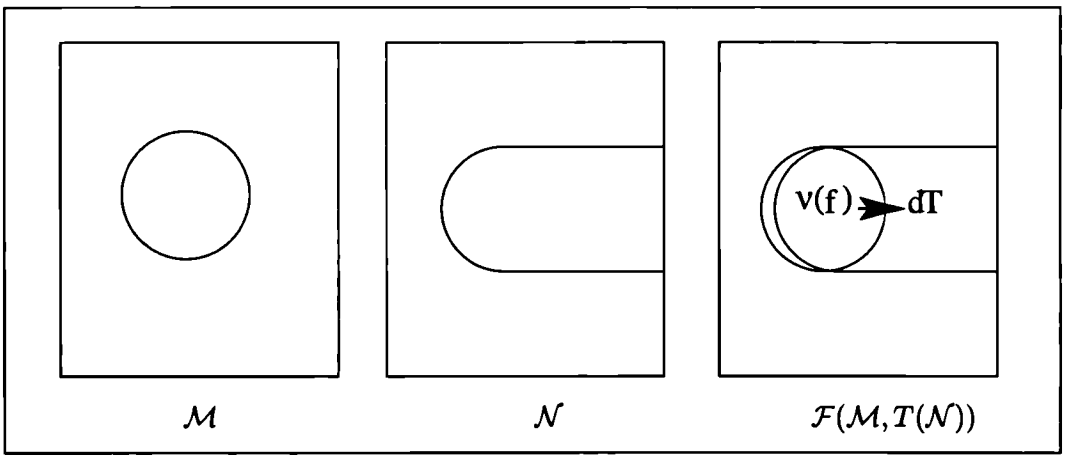


Figure 3.5: An illustration of a more subtle point of alignment between two regions in different modalities (left and centre) when they are combined (right).

completely, since there will be a set of transformations for which the sizes of overlapping regions remain unchanged. It is also possible to consider the more complex case of the behaviour of combinations of regions. The overall alignment may then be constrained in separate directions by different regions as illustrated in Figure 3.6.

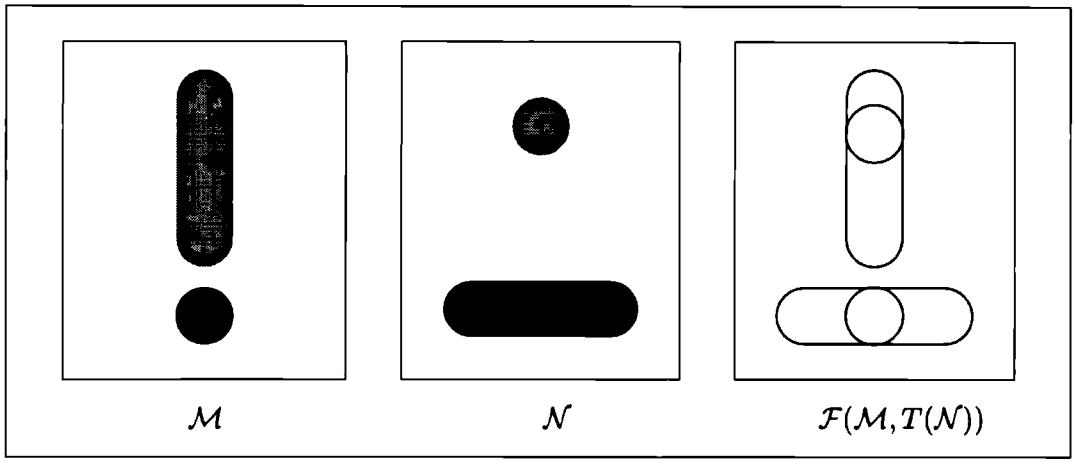


Figure 3.6: A point of 'alignment' resulting from a combination of regions

Those regions which are delineated completely by one (3.10) or both (3.9) properties will be at a maximum volume at registration. Alternatively, the case where the volume of a region reaches a minimum at registration, will be a region which will not be delineated completely by either property (3.11).

For properties which delineate identical regions (i.e. $k = l$) where the mapping between values $\tilde{M} \mapsto \tilde{N}$ in the modalities is one-to-one, the volume of all intersecting regions will either tend to zero or tend to a maximum at alignment. The resulting matrix $U(T_R)$ can then a broken up into diagonal matrices for example:

$$U(T_R) = \begin{bmatrix} 0 & \nu(k_1 \cap T_R(l_2)) & 0 & 0 \\ \nu(k_2 \cap T_R(l_1)) & 0 & 0 & 0 \\ 0 & 0 & \nu(k_3 \cap T_R(l_3)) & 0 \\ 0 & 0 & 0 & \nu(k_4 \cap T_R(l_4)) \end{bmatrix}. \quad (3.20)$$

The special case of this is when the ordering of values in the two modalities $M \mapsto N$ is such that the whole matrix $U(T_R)$ is diagonal,

$$U(T_R) = \begin{bmatrix} \nu(k_1 \cap T_R(l_1)) & 0 & 0 & 0 \\ 0 & \nu(k_2 \cap T_R(l_2)) & 0 & 0 \\ 0 & 0 & \nu(k_3 \cap T_R(l_3)) & 0 \\ 0 & 0 & 0 & \nu(k_4 \cap T_R(l_4)) \end{bmatrix}. \quad (3.21)$$

In this case there is a linear relationship between values in the two images and measures such as correlation coefficient may be used as an indication of alignment.

For modalities which delineate different regions, there may be examples of all types of region volume behaviour as the images approach true alignment. In particular there may be some regions which neither reach a minimum or maximum volume at registration. If we cannot assume correspondence of any of the regions, we don't know how the intersecting volume of each pair of regions should behave. It is interesting to note that

this problem can be related to the 'softassign' Procrustes point matching problem. An approach to this was recently proposed by Rangarajan *et al.* [75], aligning points for which there is no prior knowledge of correspondence.

3.2.5 Statistics: A Global View of Region Alignment

In the description of the problem so far we have looked at the behaviour of *individual* regions only. An alternative view is to look at the behaviour of volumes of regions $U(T)$ as a whole. We can do this by first developing a statistical view of their behaviour.

From Regions to Probabilities

A representation which captures the relationships between volumes of all regions delineated by a pair of images is the joint histogram of values. For each pair of values of two properties, the number of points at which those values occur together in the imaged volume of overlap is recorded in a 2 dimensional array, resulting in a direct estimate of the matrix U (3.14). By normalising each volume by the total imaged volume of overlap,

$$\nu_O(T) = \nu\left(V_M \cap T(V_N)\right) \tag{3.22}$$

we can form a *joint probability distribution*,

$$P(\tilde{M}, \tilde{N}) = \frac{1}{\nu_O(T)} U(T). \tag{3.23}$$

Here, individual elements,

$$P(\tilde{M}, \tilde{N}) = \begin{bmatrix} p\{\tilde{m}_1, \tilde{n}_1\} & p\{\tilde{m}_1, \tilde{n}_2\} & . & . & p\{\tilde{m}_1, \tilde{n}_j\} \\ p\{\tilde{m}_2, \tilde{n}_1\} & p\{\tilde{m}_2, \tilde{n}_2\} & . & . & p\{\tilde{m}_2, \tilde{n}_j\} \\ . & . & . & . & . \\ p\{\tilde{m}_i, \tilde{n}_1\} & p\{\tilde{m}_i, \tilde{n}_2\} & . & . & p\{\tilde{m}_i, \tilde{n}_j\} \end{bmatrix}. \tag{3.24}$$

represent the probability of occurrence of individual pairs of values in the imaged region[1].

From this we can extract the marginal probability distributions of the sets of values in each view of the scene alone,

$$p\{\tilde{m}\} = \sum_{\tilde{n} \in \tilde{N}} p\{\tilde{m}, \tilde{n}\}, \tag{3.25}$$

and,

$$p\{\tilde{n}\} = \sum_{\tilde{m} \in \tilde{M}} p\{\tilde{m}, \tilde{n}\}. \tag{3.26}$$

[1] In the strict sense these are not really probabilities but 'memberships', as there is a definite fraction of locations exhibiting a particular pair of values, and not a 'probability' of a voxel having a particular value. This leads to a fuzzy set theory view of the problem.

3.3 Early Approaches

3.3.1 Aligning Images of the Same View

Sum of Absolute Value Difference (SAVD)

The simplest and most direct measure of similarity of two image values is given by their absolute difference. An overall measure of alignment may then be determined by the average absolute difference over each of the Q measurement points,

$$SAVD(T) = \frac{1}{Q}\sum_{i=1}^{Q} \|\tilde{m}(\mathbf{x}_i) - \tilde{n}(T(\mathbf{x}_i))\| \tag{3.27}$$

In applying this measure we are assuming that the image values are effectively calibrated to the same scale so that corresponding objects exhibit the same measurement value. If this is the case then SAVD will tend to zero at alignment, if this is not the case then the behaviour is less predictable, depending on the measurement values delineating the largest regions. It is interesting to consider the amount of influence that a boundary between two regions has on the overall measure for two transformations say T_1 and T_2. Equation 3.27 shows that it will obviously depend on the difference of values across the boundary $\tilde{n}(T_1(x_i)) - \tilde{n}(T_2(x_i))$. The overall influence is dependent on the length and orientation of the boundary. The influence though does not depend on the size of the objects at either side of the boundary. High contrast boundaries between small objects may therefore have a greater influence on the measure than low contrast boundaries between large objects.

This measure is ideal in cases where two images are identical except for noise. Example applications include the alignment of images acquired to investigate the progression or regression of disease, or changes in the chemical state in the brain. Application of this approach to aligning other modalities where image values are not directly related is limited. Even the simplest application of matching one MR image to another of the same sequence is complicated by the varied scaling between two seemingly identical acquisitions.

Correlation

In the simplest case of within modality registration, for example aligning two MR images of the same sequence, registration results in a strong linear relationship between corresponding values in the two images. Misalignment then breaks down this relationship as different values over-lie each other. In this case an obvious measure of similarity $S(.)$ would be one which determines the fit of a line to the distribution of corresponding

values. Such a measure is correlation. In deriving a match from the fit of the data to a line, we can avoid the problems of intensity scaling experienced with the SAVD measure.

This approach is employed in the most commonly used measure of image alignment, correlation, and is expressed simply as the sum of the product of all the pairs of values in the images,

$$\gamma(T) = \sum_{i=1}^{Q} \tilde{m}(\mathbf{x}_i).\tilde{n}(T(\mathbf{x}_i)) \tag{3.28}$$

There are two basic limitations of this function when used as a measure of image alignment. Firstly, it is not independent of the number of points over which it is evaluated. In comparing two transformations, there may be significantly different regions of overlap between the imaged volumes, and so appreciably different numbers of points Q available for evaluation of γ. In such a case the correlation would tend to be greater for the transformation resulting in the greater volume of overlap. This can be avoided simply by dividing γ by the number of points,

$$\tilde{\gamma}(T) = \frac{\gamma(T)}{Q} \tag{3.29}$$

to provide an estimate of the *normalised cross correlation.* A further problem is that $\tilde{\gamma}$ is also not independent of the overall level of the signal in the region of overlap and so the measure would tend to favour alignments which include larger image values. As a result, a better measure of alignment of finite images is the *Correlation Coefficient* given by,

$$\Gamma(T) = \frac{\sum_{i=1}^{Q}(\tilde{m}(\mathbf{x}_i) - \overline{m}).(\tilde{n}(T(\mathbf{x}_i)) - \overline{n})}{\{\sum_{i=1}^{Q}(\tilde{m}(\mathbf{x}_i) - \overline{m})^2.\sum_{i=1}^{Q}(\tilde{n}(T(\mathbf{x}_i)) - \overline{n})^2\}^{1/2}}, \tag{3.30}$$

where $\overline{m}$ and $\overline{n}$ are the mean values of the two images over the set of Q points. Mathematically the correlation coefficient is a measure of the residual errors from the fitting of a line to the data by minimisation of least squares. In the field of signal processing the correlation of a signal with a known signal is *template matching. Matched filtering* or convolution [43] is simply the correlation of a signal with a known signal reflected along each of its axes. Correlation would be expected to be useful only for image types where the relationship between values in the two images is predominantly linear at registration.

Some of the earliest work on approaches to image alignment were in the area of satellite image registration. Here 2D image matching was carried out both between images of the same type and images of different types. Anuta [1] investigated the use of the newly developed FFT algorithms to carry out sub-pixel alignment by correlation over regions of the two images. He applied this to the registration of mutli-spectral infra-red imagery acquired at different times from aircraft and to the alignment of images acquired by Apollo 9. In order to deal with the problems of different image content, pre-processing

steps were investigated to improve the registration performance. He examined the use of thresholded gradient operations to locate edges, and intensity clustering to classify and label images before matching. He showed little improvement was provided by the use of these techniques and in particular, showed how the performance degraded for all approaches when matching an infra-red image with others of increasingly different wavelength. It is interesting to note that the main application of the registered imagery was then multi-spectral clustering and classification.

Svedlow *et al.* [88] looked at the problem of aligning different infra-red channels of Landsat imagery. In particular they examined the use of the Correlation, Correlation Coefficient and Sum of Absolute Value Differences (SAVD) in conjunction with different pre-processing steps.

Yao and co-workers [109] examined the problem of aligning Landsat Thematic Mapper imagery to Seasat Synthetic Aperture Radar imagery. They employed normalised Cross Correlation derived from a number of individual rectangular regions in the images. Estimates derived from each block were then sorted to remove outliers and then combined. They found that edge processing of the imagery provided appreciable improvements in performance when aligning with radar imagery.

In medical applications, Apicella *et al.* [2] investigated the use of Fourier correlation of MR and PET images of the brain. Employing a Fourier decoupling of the rotation and translation components of alignment, an efficient direct estimate of registration was proposed. The results were limited to 2D slice registration on one dataset.

Variance of Intensity Ratio

This is a simple statistical measure proposed by Woods et al [107] for aligning one PET image to another of the same patient. Empirically, when the pair of images are correctly aligned then for each pair of corresponding values in the images, the ratio,

$$r_i(T) = \frac{\tilde{m}(\mathbf{x_i})}{\tilde{n}(T(\mathbf{x_i}))}. \tag{3.31}$$

of their values is evaluated. The standard deviation of this ratio σ_r is evaluated over all i voxel pairs. This is then normalised by dividing by the mean value $\overline{r}$ of each of the voxel pairs,

$$\mathcal{V} = \frac{\sigma_r}{\overline{r}}, \tag{3.32}$$

giving a final measure of alignment which is minimised at registration. This is closely related to the correlation coefficient and assumes that there is a constant scaling factor relating the values in one image to the other at registration.

3.3.2 Enforcing Similarity:Correlation of Pre-Processed Images

Correlation of Remapped Image Values

One approach to evaluating similarity between measurements of different properties is to try and enforce a simple relationship between values prior to evaluation of alignment. This can be done by applying a remapping, say $\mathcal{I}(\tilde{m})$ and $\mathcal{J}(\tilde{n})$, of values within the images. A similarity measure,

$$\mathcal{S}(T) = \sum_{i=1}^{Q} S(\mathcal{I}(\tilde{m}(\mathbf{x}_i)), \mathcal{J}(\tilde{n}(T(\mathbf{x}_i)))), \tag{3.33}$$

can then be derived from these remapped values. If the remapping provides adequate linearity in the relationship, for example:

$$\begin{bmatrix} 0 & \nu(k_1 \cap T_R(l_2)) & 0 & 0 \\ \nu(k_2 \cap T_R(l_1)) & 0 & 0 & 0 \\ 0 & 0 & \nu(k_3 \cap T_R(l_3)) & 0 \\ 0 & 0 & 0 & \nu(k_4 \cap T_R(l_4)) \end{bmatrix} \tag{3.34}$$

$$\mapsto \begin{bmatrix} \nu(k_1 \cup k_2 \cap T_R(l_1 \cup l_2)) & 0 & 0 \\ 0 & \nu(k_3 \cap T_R(l_3)) & 0 \\ 0 & 0 & \nu(k_4 \cap T_R(l_4)) \end{bmatrix}, \tag{3.35}$$

at registration (here combining the first two rows and columns), then maximisation of correlation can then be used to align the images.

Van den Elsen *et al.* [95] employed a simple remapping of CT values to increase the linearity of the relationship between CT and MR values, enabling correlation to be used as a measure of alignment. This made use of the strong differentiation between soft tissue and bone regions in CT. Two approaches were described. The first approach was devised to enable the use of negative correlation to align bone features in the two images. This was achieved by applying a simple 'delayed ramp' mapping to the CT values c above a CT bone threshold C_b,

$$\mathcal{J}(c) = \begin{cases} 0 & c < C_b \\ (c - C_b).K_b & \text{otherwise} \end{cases}, \tag{3.36}$$

where K_b determines the gradient of the ramp. This produces a 'positive' bone image from CT, which can then be registered with the 'negative' bone image in a conventional MR sequence by minimising correlation. The approach employed a multi-start, multi-resolution hierarchical optimisation scheme to locate a global optimum. This was perhaps because the large number of soft tissue structures in the MR image, which are not present in the remapped CT, create many false optima and a narrow global optima.

The second scheme was designed to use the soft tissue features for registration by again identifying them from the CT values. All values outside the soft tissue range

($C_{s1} \approx 500$ to $C_{s3} \approx 1900$) were set to zero, and within that range a triangular remapping was applied upto a value ($C_{s2} \approx 700$),

$$\mathcal{J}(c) = \begin{cases} 0 & c \leq C_{s1} \\ (c - C_{s1}).K_s & C_{s1} < c \leq C_{s2} \\ (C_{s3} - c).K_s & c > C_{s3} \\ 0 & \text{otherwise} \end{cases}, \tag{3.37}$$

where K_s determines the gradient of the triangular remapping. This provides a 'positive' soft tissue region which can be correlated with the MR image.

Both of these approaches, use knowledge about the tissues delineated by different values in a CT image to provide a useable registration parameter space. An important question about these approaches is how the intensity remapping affects the location of boundaries between the different objects, particularly when there are partial volume effects between voxels lying on object boundaries.

Correlation of Geometrical Features

An alternative to remapping the *values* within images to allow the use of correlation, is to 'remap' or enhance the geometric structure within the images by extracting or detecting particular features. The aim here is to choose types of geometric structure present in both images and that occur at the same spatial location in the two modalities. This is related to the early work on the application of edge detection filters as a pre-processing step to enable the correlation of infra-red and radar imagery[109].

In the medical application the work of van den Elsen *et al.* [93, 94] has investigated the use of different geometrical feature detectors. In particular they looked at the use of rotation invariant geometrical ridge detectors [65] in three dimensions to detect the bone structure of the skull in both MR and CT images. Their work also made much use of scale space theory to constrain the size or 'scale' of geometric features being detected. In later work they went on to compare this approach to that based on much simpler edge detection operations.

One of the possible limitations of this approach, is the nature of the parameter space produced by the ridge and boundary detectors. In using a correlation approach rather than a true distance measure between features, the recovery of larger scale misalignments is complex. Empirically correlation is still a measure of region overlap, but these geometric feature 'regions' occupy a small portion of the field of view. Correlation of these small structures then provides a narrow optima which is expensive to locate. This is illustrated by the the plots shown in Figure 3.7. Here on the left, the upper two traces represent a region of material delineated in both modalities, and the lower trace the resulting correlation as the transformation between the two is varied round registration.

The right hand trace then represents a simplistic 'edge' image of the same scene in the two modalities. The resulting peak in the correlation of these two traces is much narrower and therefore more difficult to find. This problem is reflected in the approach to

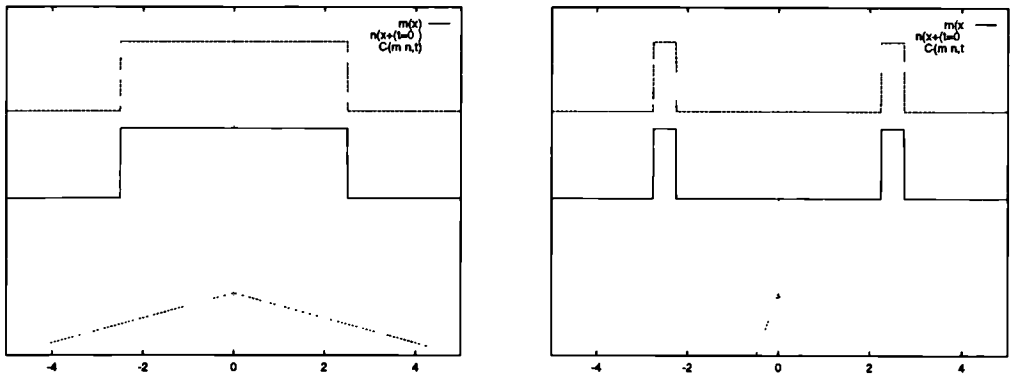


Figure 3.7: A simplified view of region (left) and boundary (right) correlation. The use of sharp edge features leads to a much narrower response to misalignment at a given scale.

optimisation which employs an hierarchical multi-resolution technique. Multiple optima are evaluated and optimised separately at increasing image resolution to locate the best overall transformation.

3.4 Statistical Measures between Different Scenes

A similarity $S(.)$ can be derived from the global relationship between voxel values in the two modalities. The problem with criteria such as correlation is that they assume some simple mathematical form (e.g. linear) for the distribution at registration. What we want is to make less constrained assumptions about the relationship between the two properties which will allow us to define other more generic measures of registration.

An alternative view of alignment is to say that two values m and n are related or *similar* simply if there are many other examples of those values occurring together in the overlapping imaged volume. Conversely if there are few examples of values m and n occurring together they can be said to be dissimilar. This is the basic idea behind a class of more generic statistical measures which look only at the occurrence of image values and not at the values themselves. Effectively we can say that in aligning the images we wish to change the transformation T in order to increase the *clustering* of the joint probability distribution so that there are fewer low probability pairs and more higher probability pairs. In terms of the delineated regions $f \in \mathcal{F}(T)$, we can seek a transformation which maximises the volume of the *majority* of the intersecting regions.

3.4.1 Corresponding Intensity Variance

Woods proposed the minimisation of corresponding variance [108], as an adaption of equation (3.32) for multi-modality registration of MR and PET imagery of the brain. Given the assumption that there will be a predominantly one-to-one mapping of values in the images at registration, then for all points in an image with a given value, the variance of all corresponding values in the other modality will be minimised at registration. We can work out for all the voxels in image $\tilde{m}(\mathbf{x})$, say with value m, the mean value of the corresponding voxels in the second image $\tilde{n}(T(\mathbf{x}))$, say $\overline{\tilde{n}}_m$. Similarly we can find the standard deviation of those values $\sigma_n(\tilde{m})$. For a particular value $\tilde{m}$ then we can work out the normalised standard deviation,

$$\sigma'_n(\tilde{m}) = \sigma_n(\tilde{m})/\overline{\tilde{n}}_m. \tag{3.38}$$

The standard deviation of the distribution of values $\tilde{n}$ for each value $\tilde{m}$ should be minimised at registration. Given a probability of occurrence for each value $\tilde{m}$, say $p(\tilde{m})$, a weighted sum of the normalised standard deviations of the PET values $\tilde{n} \in \tilde{N}$ corresponding to each MR value $\tilde{m} \in \tilde{M}$,

$$\sigma(\tilde{N}) = \sum_{\tilde{m} \in \tilde{M}} p(\tilde{m})\sigma'_n(\tilde{m}), \tag{3.39}$$

provides a measure of alignment. The weighting ensures that the measure is influenced most strongly by PET intensity variation for the most common MR values.

The basic assumption this approach makes is that, at registration, uniform regions in one image map to uniform regions in the other. As the images move away from alignment, measuring the spread of corresponding values by their variance gives a direct measure of alignment. Effectively, this is assuming that, at alignment, there is a one to one mapping of values in the images. If this is exactly true then only one modality is required and we have the case described by equation 3.10. In practice this though may be approximately true for images where there are only small differences.

If for two images at registration, one value maps to two significantly different values in the other modality, then a measure of the clustering around the mean value will provide a poor indication of alignment. The degree to which there is a direct one to one mapping between values in the two images will determine the applicability of this measure. In order to ensure a close to one-to-one mapping of values in his application, Woods [108] employed a simple MR segmentation to exclude non-brain regions, particularly scalp, from the measure.

What is also important is the *direction* of the mapping, i.e. choosing to minimise corresponding variance in the first or second modality can make a significant difference.

One modality, say MR, may delineate significantly more regions than say PET and so every value in the MR may map to only one value in the PET (except perhaps a small region of physiological abnormality), while the reverse is not true.

3.4.2 Moments of the Joint Probability Distribution

This approach was first proposed by Hill *et al* [49] from visual examination of the effects of misregistration on the feature space. Empirically, as the images approach registration, the peaks in the joint probability distribution increase in height and the regions which contain lower counts decrease in height. The registration process is re-arranging the voxels so that they occur with their most probable corresponding value in the other image, given the constraints of the image structure and the transformation.

One approach to quantifying this shift from more low probabilities in $p\{\tilde{m},\tilde{n}\}$ to a smaller number of higher probabilities in $p\{\tilde{m},\tilde{n}\}$ is to measure *skewness* in the *distribution of probabilities* in $p\{\tilde{m},\tilde{n}\}$. Common measures of skewness of a distribution are provided by the higher order moments of the distribution. Given $p\{\tilde{m},\tilde{n}\}$ we can calculate the number of occurrences of a particular probability p, say $\eta(p)$. The moment of order i of this distribution can then be evaluated:

$$\vartheta_i(\tilde{M},\tilde{N}) = \int_0^1 \eta(p).p^i dp.$$

This can be normalised by dividing by the zero moment or mass,

$$\vartheta_0(\tilde{M},\tilde{N}) = \int_0^1 \eta(p) dp.$$

Hill proposed the use of the third order moment as a measure of MR and CT alignment which is simply:

$$\vartheta'_3(\tilde{M},\tilde{N}) = \frac{\vartheta_3(\tilde{M},\tilde{N})}{\vartheta_0(\tilde{M},\tilde{N})}. \tag{3.40}$$

This has been compared experimentally to the joint entropy measure[87] and found to provide similar response to alignment of MR and PET images of the brain. Maes *et al* [64] has proposed that this is mathematically related to the information theory measures which will be described in the next chapter.

3.4.3 Clustering and Labelling

Chiron and Bizais [41] proposed a measure of alignment based on a matched filtering of a joint histogram of image values. The matched filter consisted of a Gaussian model of the noise distribution derived from the distribution of values in an example registered region of the images.

The work of Collignon [22] *et al* was the first to describe an approach to quantifying registration by applying conventional multi-spectral clustering techniques. They investigated the use of both supervised and unsupervised clustering of the joint probability distribution using fuzzy k-means clustering [38]. Following clustering they employed a maximum likelihood ([28], page 250) voxel labelling scheme from which two measures of image alignment were evaluated. Firstly they defined a Geometric Mean Maximum Registration Probability. Each label cluster is modeled as a normal distribution and the probability that a voxel value pair belongs to a cluster is calculated. The most probable cluster is chosen for each voxel pair and the product of the probabilities of each voxel value belonging to its cluster is calculated for a given transformation. In other words: for each pair of values occurring together, a probability for those values occurring together is determined by its distance from the nearest cluster. This product should then be maximised at registration.

A second measure was derived from a definition of unlabeled voxels as a threshold of the Mahalanobis distance to the nearest cluster centre. A simple count of the number of unlabeled voxels in the distribution could then be minimised to achieve alignment. Their experimental results illustrated the behaviour of the measure with respect to misregistration, but they did not go on to use it in recovering alignment. In particular they did not investigate the problems in achieving an effective labelling from a poor initial starting estimate where the joint probability distribution contains unwanted but distinct clusters.

3.4.4 Describing and Optimising the Intensity Transformation

Rather than looking at the statistical relationship between image values, it is possible to treat the relationship between intensities as a transformation. This intensity transformation function can then be estimated and optimised in the same way as the geometric transformation which describes alignment. An example of such an approach is that of Friston *et. al.* [37]. Here both the intensity and geometric transformations are approximated by Taylor series and then solved directly. This approach was applied to PET-PET and MR-PET brain image alignment but was not illustrated for truncated images or MR-CT image pairs. In such cases the more complex intensity transformation coupled with the limited and changing volume of overlap may mean that a Taylor series solution of the geometric and intensity transformation is not valid for many MR-CT applications.

A related approach is that of Feldmar *et al.* [34]. This expresses the 3D registration in terms of the alignment of 4D points, made up of a 3D location and the intensity

value. Registration is then achieved by applying an extension of the iterative closest point algorithm described for surface matching in chapter 2. The key factor in this approach is the relationship between the geometric and intensity components of the minimisation. It remains to be shown whether approaches such as these are robust to a range of clinical data and modality combinations.

3.5 Summary

This chapter began by looking at image alignment in terms of overlapping regions of material which are delineated by values in two modalities. As images are brought into alignment, the volume of overlap of these regions can give an indication of alignment. By seeking a maximum or minimum overlap of specific regions it is possible to define an optimal point of alignment. This provides a similar approach to that described in the previous chapter, but again requires knowledge of correspondence, this time between particular values in the two images.

The overlap of all delineated regions can be summarised by the joint probability distribution of corresponding image values. Without knowledge of specific region correspondence it is possible to define a *global* measure of alignment from the behaviour of regions as a whole.

Early work on multi-modality alignment began with the use of correlation and absolute difference measures between pre-processed imagery. In some cases, by extracting or enhancing similar features in the two modalities it is possible to impose a simple relationship between image values. In medical image registration, preprocessing of both intensity and geometric forms have been employed to increase the similarity of images acquired by different modalities.

More generic statistical approaches began to be developed which did not assume a predominantly linear relationship between values and avoid the use of a heuristic preprocessing stage. The important step at this point was to begin to define registration measures based only on the co-occurrence of values and not the values themselves. Effectively we can then infer similarity $\mathcal{S}(m,n)$ of a pair of values simply by observing other examples of them occurring together in the imaged space. Registration can then be viewed as seeking a transformation which makes more of those pairs of values occur together.

Chapter 4

Information Theory and Image Alignment

4.1 Introduction

This chapter is concerned with development of entropy based measures of multi-modality alignment. As with the previous chapter we are essentially dealing only with views of the scene provided by different modalities, and not with the characteristics of the images themselves. As a result we deal with the actual values of materials within a scene $\tilde{M}$ and $\tilde{N}$. Entropy and related measures are derived from the probability of occurrence of these values. Throughout this chapter we are dealing with probabilities, but it is important to remember what these probabilities represent: the proportion of material exhibiting a given value in the scene.

The chapter begins by introducing some of the fundamental ideas behind the measurement of 'information' and how this can be applied to image data. Information theory is then used as a tool in deriving measures of alignment between sets of values in a pair of views of a scene. Joint entropy derived from the joint probability distribution provides the starting point for relating the information content of a pair of images. Its properties are examined as a measure of image alignment and from this a set of related measures including mutual information and normalised mutual information are developed and compared.

The final section examines the behaviour of the measures, in particular their normalisation with respect to the data from which they are derived. A simple simulation illustrates the difference in the behaviour of many of the statistical and information theoretic measures as the field of view of a scene is varied.

4.2 Information and Entropy

Before going on to describe plausible measures of alignment, we can first cover some of the background ideas dealing with the application of information theory to image data. One approach to developing ideas about information theory is based on the concept of uncertainty.

4.2.1 Information from Uncertainty

If we examine the probabilities of values occurring in an image, it is common for there to be a range of probabilities, since some values are rare and some are common. In predicting what value a voxel has we can form an estimate of the *uncertainty* of our guess at the value given the observed distribution of probabilities. This has two extremes:

- If all probabilities are equal, uncertainty in guessing what value a given voxel might have is largest.
- If there is only one value in the scene then, the uncertainty in guessing the value of a given voxel would be zero.

If we learn the value of a measurement that we were very uncertain about (i.e. it would be difficult to guess) then we gain a large amount of information. Conversely if we learn a value that would be easy to guess because it has a high probability of occurrence, then we only gain a small amount of information. If we are given a set of values we can then look at the average amount of information provided by the set of values. Essentially we can say an image with similar amounts of all possible values contains more information than an image where the majority of voxels have the same value. In order to express this mathematically a number of axioms have been proposed to describe how we would like a possible information measure to behave. Given an information measure $H(M)$ derived from the probabilities of occurrence of i possible values, some of the requirements that have been proposed for a measure $H(p(\tilde{m}_1), p(\tilde{m}_2), ..., p(\tilde{m}_i))$ of information include ([76] page 80):

Continuity Small changes in probabilities should give only small changes in the overall information content.

Symmetry The information content should be independent of ordering of probabilities.

$$H(p_1.p_2....,p_i) = H(p_2.p_1,...,p_i)$$

Extremal Property When all probabilities are equal the average uncertainty and so information must be maximised:

$$\text{Maximum of } H(p_1, p_2, ..., p_i) = H\left(\frac{1}{i}, \frac{1}{i}, ..., \frac{1}{i}\right)$$

Additivity Combining information from subsets: Given probabilities of a set of values with information content,

$$H_A = H(p_1, \ldots, p_{i-1}, p_i)$$

and that say one of these values (p_i) can be further divided into a second set, each with their own probability of occurrence $\{q_1 \ldots q_j\}$ and information content,

$$H_B = H\left(\frac{q_1}{p_i}, \ldots, \frac{q_j}{p_i}\right)$$

then the information provided by all the symbols should be expressed as

$$H_O = H(p_1, \ldots, p_{i-1}, q_1, \ldots, q_j) = H_A + p_i H_B.$$

4.2.2 A Measure of Information

There have been many plausible functions proposed for $H()$. An early example of an attempt at defining a measure of information is that of Fischer [35]. The most commonly used measure of information which satisfies these requirements though is the Shannon-Wiener entropy measure [76]. Here the average information supplied by a set of i symbols whose probabilities are given by $\{p_1, p_2, \ldots p_i\}$, can be expressed as,

$$H(p_1, p_2, ..., p_i) = -\sum_{s=1}^{i} p_s log p_s. \tag{4.1}$$

This forms the basis for a large body of work originally termed 'communication theory' developed to study the performance of communication systems.

In terms of image processing the measure can be applied directly to the occurrence of the values in an image $H()$, to form a *first order* estimate of the entropy of the image. Higher orders can be derived from the occurrence of sets of values together in an image.

The concept of the entropy of a set of data or an image has been applied to a number of image processing tasks. One of the most successful has been the maximum entropy [84] techniques of image reconstruction. This approach is derived from a model of the overall probability of a particular image, made up of pixels into which fall 'counts' to form intensities at points in the image. Given a set of data and an algorithm to reconstruct an image from it, there are a number of parameters which determine the final form of the image. Maximum entropy techniques evaluate the entropy of the final image and effectively search for the most probable individual image for the given data.

4.2.3 Information Measures from Two Views of a Scene

Given a pair of views of a scene $\tilde{m}(\mathbf{x})$ and $\tilde{n}(\mathbf{x})$ provided by different modalities, with sets of values $\tilde{M}$ and $\tilde{N}$ and marginal probabilities as in (3.25) and (3.26), the average information provided by the values in each scene (marginal entropy) can be evaluated:

$$H(\tilde{M}) = -\sum_{\tilde{m}\in\tilde{M}} p\{\tilde{m}\}\log(p\{\tilde{m}\}) \tag{4.2}$$

$$H(\tilde{N}) = -\sum_{\tilde{n}\in\tilde{N}} p\{\tilde{n}\}\log(p\{\tilde{n}\}) \tag{4.3}$$

If we return to the region based view of the alignment problem described in section 3.2, it is possible to relate entropy to the size (probability) of regions delineated within the two views. If a given value $\tilde{m}$ in a modality delineates a region of the material then we can relate the probability of that value $p\{\tilde{m}\}$ directly to the volume of that region in the image divided by the total imaged volume. The simplest case is where we have a binary scene with foreground and background intensities. This can be equated to a binary symmetric channel as illustrated in Figure 4.1. Information content is maximised when the area of the two regions are the same.

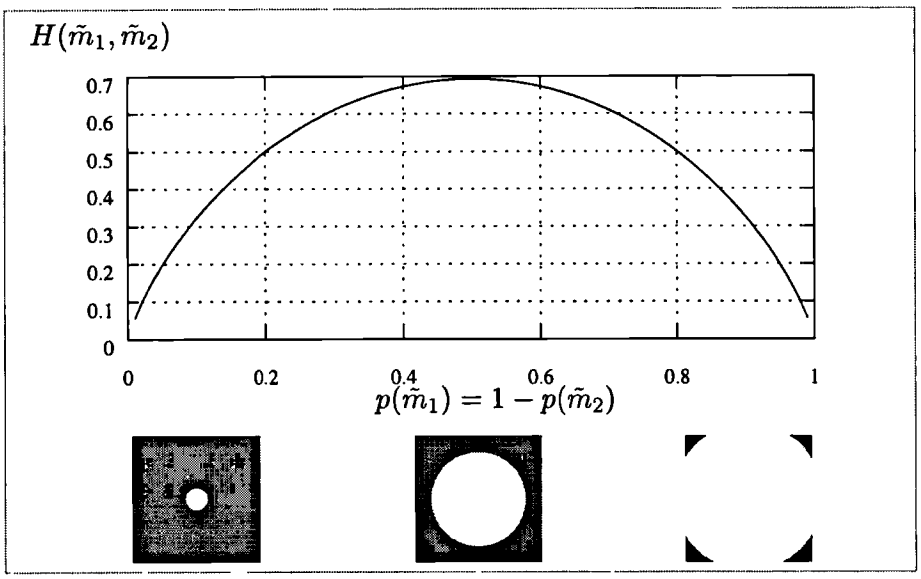


Figure 4.1: The information content of a binary symmetric channel: This can be equated to the information content of a binary image consisting of foreground ($\tilde{m}_1$) and background ($\tilde{m}_2$) where $p(\tilde{m}_1) = 1 - p(\tilde{m}_2)$. The horizontal axis in this graph therefore corresponds to varying the ratio of background to foreground area in the images of the white circle shown below.

For a more realistic scene, if we have a large number of different anatomical regions delineated by distinct values, then the entropy is derived from the probability of each of

those values. The more similar the probabilities, the greater the average entropy of the image, as illustrated in Figure 4.2.

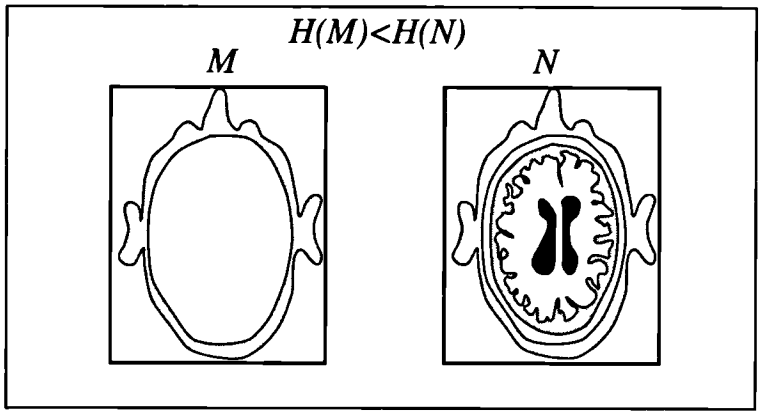


Figure 4.2: The first order entropy of the values in an image increases when a modality delineates more regions of different tissue.

Joint Entropy

When two views of a scene are combined by a transformation mapping from points in one view to the other, the joint probability distribution (equation 3.23) tells us how often pairs of values occur together. The information content of this combined scene can be evaluated in a similar way to that of a scalar scene by forming an estimate of the joint entropy from the joint probability distribution,

$$H(\tilde{M},\tilde{N}) = -\sum_{\tilde{n}\in\tilde{N}}\sum_{\tilde{m}\in\tilde{M}} p\{\tilde{m},\tilde{n}\}\log(p\{\tilde{m},\tilde{n}\}). \tag{4.4}$$

As discussed earlier, the values of $p\{\tilde{m},\tilde{n}\}$ are directly related to the overlap of the regions delineated by the values $\tilde{m}$ and $\tilde{n}$ together in the imaged volume. In terms of communication theory, $\tilde{M}$ and $\tilde{N}$ are the *sample spaces* of the two modalities and the combined image is termed the *product space* $\tilde{M}\otimes\tilde{N}$.

4.3 Information Measures and Alignment

4.3.1 Joint Entropy and Image Alignment

The use of joint entropy as a measure of multi-modality image alignment has been proposed independently in two recent papers [21, 87]. In aligning different modalities we wish to use any shared features in the two images to define the registration. If structures are shared between the two images and the images are misaligned, then in the combined image these structures will be duplicated. For example, when a transaxial slice through the head is mis-aligned, there may be four eyes and four ears. As the images are

brought into alignment the duplication of features is reduced and the combined image is simplified, as shown in Figure 4.3.

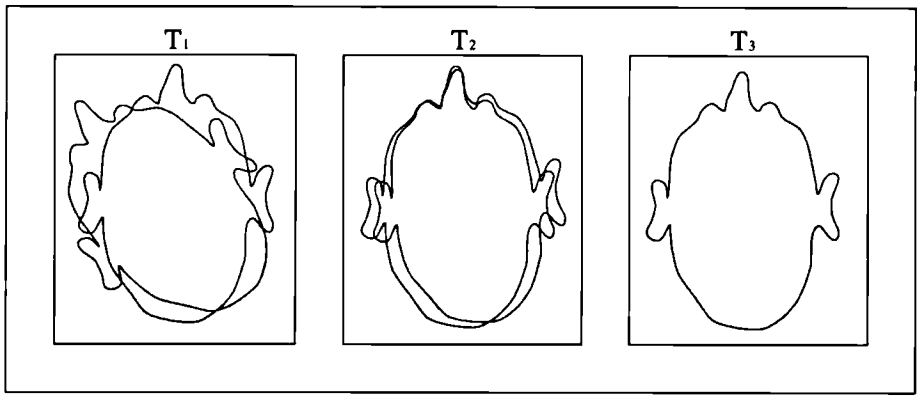


Figure 4.3: Sketch illustrating the removal of 'duplicated' regions in the combined image by alignment of shared features.

We can think of the 'combined' image, where pairs of values occur together, as a single valued image where each different value corresponds to a particular pair of values occurring in the two source images. We may then express the joint probability distribution as a 1D vector, where each probability is simply derived from the size of each intersecting region.

If we look again at the plot for the binary image (Figure 4.1), this shows that the function $H(\tilde{m}_1, \tilde{m}_2)$ reaches a minimum when values are highly probable (larger regions) or not probable (smaller regions). By minimising joint entropy, overlapping regions will therefore be forced to either have an increased overlap or a decreased overlap. This, in effect, is exactly what we are doing when aligning any shared regions which already have a significant portion of their area overlapping. We increase the area of the larger aligned part and decrease the area of smaller mis-aligned parts.

In using joint entropy as a measure of alignment we are making one important assumption: that the large regions in the two images are to be aligned and their region of overlap should increase as the images approach registration. If there are large regions in both images, but at alignment their overlap is not maximised, then joint entropy will not be a minimum at registration. This case will discussed in more detail later in the thesis.

4.3.2 Normalisation: The Need for Relative Entropies

If we take a closer look at the use of joint entropy as a measure of alignment, we can see in practice that there is a problem. In bringing two images into alignment we are

interested in essentially comparing two alignments to decide which is better. In deriving a measure of alignment from images with limited field of view (often in the clinical case, very limited) then any measure of structure or 'information content' of the combined image will be a function of the structure in the two images in their region of overlap. By minimising joint entropy we are simply trying to find the combined image which delineates least structure, not necessarily the most corresponding structure. This problem can be related back to the derivation of the measures of correlation, normalised correlation and correlation coefficient for image matching discussed in the previous chapter.

A second and related problem can be seen by considering the larger scale behaviour of the measure as mis-alignment increases. As we continue to misregister we would hope the measure will continue to degrade, tending toward a maximum of joint entropy. Going back to the axioms of information measures, this occurs when we have identical probabilities for each of the say w pairs of measurement values, so as we misregister we assume,

$$P\{\tilde{M},\tilde{N}\} \rightarrow \begin{bmatrix} 1/w & 1/w & 1/w & 1/w & 1/w \\ 1/w & 1/w & 1/w & 1/w & 1/w \\ 1/w & 1/w & 1/w & 1/w & 1/w \\ 1/w & 1/w & 1/w & 1/w & 1/w \end{bmatrix}$$

For general medical images this is obviously not the case, and in fact as misalignment increases, tissue in one modality overlies air (say $\tilde{m}_0$ and $\tilde{n}_0$) in the other, and at the limit,

$$P\{\tilde{M},\tilde{N}\} \rightarrow \begin{bmatrix} p\{\tilde{m}_0,\tilde{n}_0\} & p\{\tilde{m}_0,\tilde{n}_1\} & p\{\tilde{m}_0,\tilde{n}_2\} & p\{\tilde{m}_0,\tilde{n}_3\} & p\{\tilde{m}_0,\tilde{n}_4\} \\ p\{m_1,\tilde{n}_0\} & 0 & 0 & 0 & 0 \\ p\{m_2,\tilde{n}_0\} & 0 & 0 & 0 & 0 \\ p\{m_3,\tilde{n}_0\} & 0 & 0 & 0 & 0 \\ p\{m_4,\tilde{n}_0\} & 0 & 0 & 0 & 0 \end{bmatrix}.$$

4.3.3 Mutual Information

What we need to do is to relate changes in the value of the joint entropy $H(\tilde{M},\tilde{N})$ back to the marginal entropies of the two images $H(\tilde{M})$ and $H(\tilde{N})$ alone. One measure which provides this is relative entropy or mutual information (or transinformation), which was proposed independently by Collignon *et al.* [21] and Viola and Wells [99] as a measure of alignment. This statistic was initially derived in communication theory as a measure of information between the transmitter and receiver at either end of a communication channel. Simply, this expresses the joint entropy with respect to the marginal entropies

$$I(\tilde{M};\tilde{N}) = H(\tilde{M}) + H(\tilde{N}) - H(\tilde{M},\tilde{N}). \tag{4.5}$$

A useful way of visualising the relationship between these entropies is provided by a Venn diagram shown in Figure 4.4. Here the size of the circles represents the value

of the particular entropy. The overlapping areas represent the shared information or relative entropies. The entropies $H(\tilde{M})$, $H(\tilde{N})$ and $H(\tilde{M},\tilde{N})$ can be derived from the

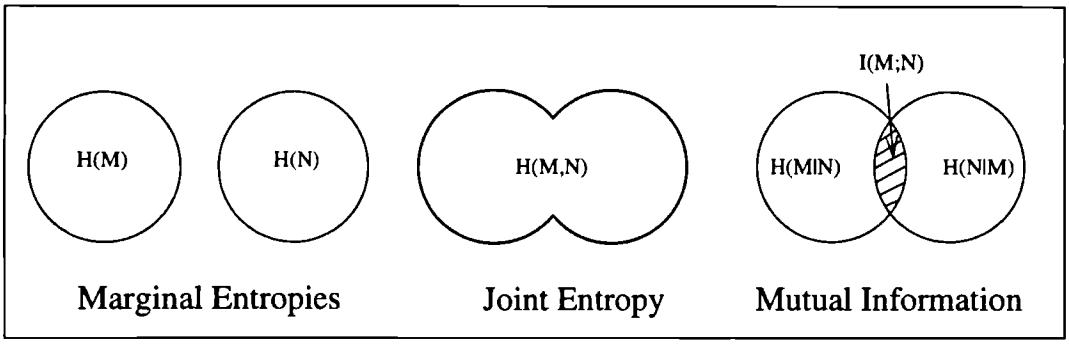


Figure 4.4: A set theory representation of the entropies involved when combining two images

probability of occurrence of values $\tilde{M}$ and $\tilde{N}$,

$$H(\tilde{M}) = \sum_{\tilde{m}\in\tilde{M}} p\{\tilde{m}\} log \frac{1}{p\{\tilde{m}\}} \tag{4.6}$$

$$H(\tilde{N}) = \sum_{\tilde{n}\in\tilde{N}} p\{\tilde{n}\} log \frac{1}{p\{\tilde{m}\}} \tag{4.7}$$

$$H(\tilde{M},\tilde{N}) = \sum_{\tilde{m}\in\tilde{M}} \sum_{\tilde{n}\in\tilde{N}} p\{\tilde{m},\tilde{n}\} log \frac{1}{p\{\tilde{m},\tilde{n}\}}. \tag{4.8}$$

Substituting these into equation (4.5), and given that $p\{\tilde{m}\} = \sum_{\tilde{n}\in\tilde{N}} p\{\tilde{m},\tilde{n}\}$ and $p\{\tilde{n}\} = \sum_{\tilde{m}\in\tilde{M}} p\{\tilde{m},\tilde{n}\}$ we have:

$$\begin{aligned} I(\tilde{M};\tilde{N}) = & \sum_{\tilde{m}\in\tilde{M}} \sum_{\tilde{n}\in\tilde{N}} p\{\tilde{m},\tilde{n}\} log \frac{1}{p\{\tilde{m}\}} + \\ & \sum_{\tilde{n}\in\tilde{N}} \sum_{\tilde{m}\in\tilde{M}} p\{\tilde{m},\tilde{n}\} log \frac{1}{p\{\tilde{n}\}} - \\ & \sum_{\tilde{m}\in\tilde{M}} \sum_{\tilde{n}\in\tilde{N}} p\{\tilde{m},\tilde{n}\} log \frac{1}{p\{\tilde{m},\tilde{n}\}}. \end{aligned} \tag{4.9}$$

which can be rearranged to give a simpler form:

$$I(\tilde{M};\tilde{N}) = \sum_{\tilde{m}\in\tilde{M}} \sum_{\tilde{n}\in\tilde{N}} p\{\tilde{m},\tilde{n}\} \log \frac{p\{\tilde{m},\tilde{n}\}}{p\{\tilde{m}\}p\{\tilde{n}\}} \tag{4.10}$$

From this representation, remembering that the probabilities are derived from volumes of overlapping regions, $I(\tilde{M};\tilde{N})$ relates the size of the overlap of a pair of regions ($p\{\tilde{m},\tilde{n}\}$), to their total size ($p\{\tilde{m}\}$ and $p\{\tilde{n}\}$).

4.3.4 Further Constraints

The fundamental problem in using mutual information for our application is that it was developed in communication theory as a direct measure of the quantity of information passed between transmitter and receiver. For any given channel this will be dependent on the information transmitted. In our application we wish to compare an information measure derived for two transformation estimates. These estimates will correspond to different overlaps of the two image volumes, and therefore the amount of transmitted (and received) information may vary between the two transformations simply because of the change in overlap. Such a case is illustrated by the Venn diagrams of Figure 4.5. Here

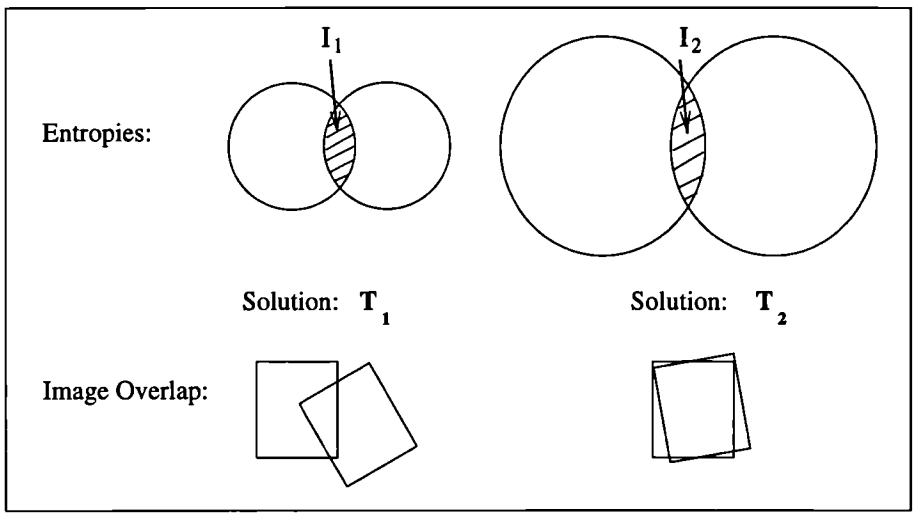


Figure 4.5: Here $I_2 > I_1$ so mutual information chooses the registration solution on the right, although the proportion of the information shared by the two images is less.

the mutual information is greater for the solution on the right, although the proportion of the information in two the images which is shared, is much less than the solution on the left. Later in the thesis example images where this occurs will be examined in more detail.

Normalising Mutual Information

In order to account for changes in the proportion of mutual information we want to look at the amount of mutual information $I(\tilde{M};\tilde{N})$, with respect to the information provided by the individual images $H(\tilde{M})$ and $H(\tilde{N})$. A direct approach to normalisation proposed in this work is to evaluate the ratio of the joint and marginal entropies

$$Y(\tilde{M};\tilde{N}) = \frac{H(\tilde{M}) + H(\tilde{N})}{H(\tilde{M},\tilde{N})}. \tag{4.11}$$

This we refer to as *Normalised Mutual Information*, and will be examined in greater detail during this thesis.

Alternative approaches to achieving the same goal have subsequently been proposed in the literature. These include the entropy correlation coefficient,

$$C_e(\tilde{M},\tilde{N}) = \frac{2I(\tilde{M};\tilde{N})}{H(\tilde{M})+H(\tilde{N})}. \tag{4.12}$$

and,

$$\rho(\tilde{M},\tilde{N}) = H(\tilde{M},\tilde{N}) - I(\tilde{M};\tilde{N}). \tag{4.13}$$

mentioned by Maes *et al.* [64].

There is a further possibly useful constraint on alignment which is illustrated in Figure 4.6. We may seek a transformation which maximises the shared information and, which also balances the amount of unshared information provided by the two images.

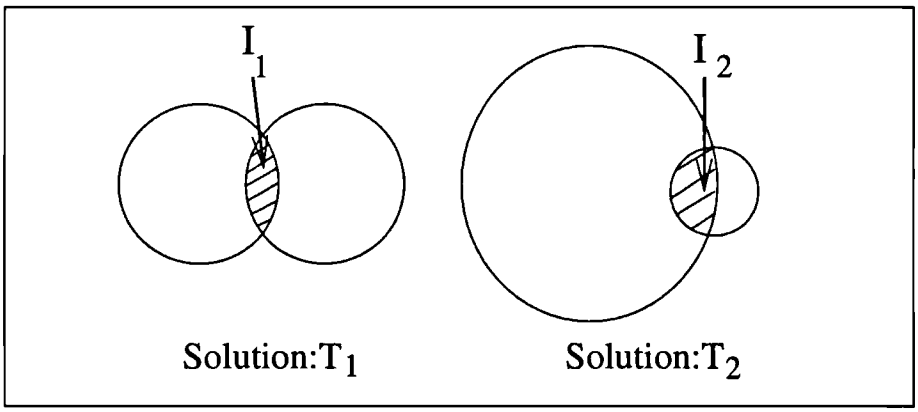


Figure 4.6: A case where mutual information does not distinguish between two plausible alignments. A better registration may be one which has a similar amount of information provided by the two images.

This is a rather more subtle constraint and requires us to introduce an assumption about the ratio of 'unshared' information supplied by the two images. This will be dependent on both the content of the images and their region of overlap. Since we are interested in generic measures applicable to all modality combinations, this approach will not be examined further in this thesis.

4.4 The Image Overlap Problem: A Simulation

If we look again at the expression for the joint probability distribution (3.23), we can see that any alignment statistic will be affected by both the alignment of regions in the two images, determined by $U(T)$, and the region of overlap of the two images $\nu_O(T)$. In addition, the volume of individual overlapping regions $U(T)$ may also be influenced by the finite extent of the images. There will be some regions $f \in \mathcal{F}$ in the field of view which will be truncated by the imaged field, as illustrated in Figure 4.7. In these

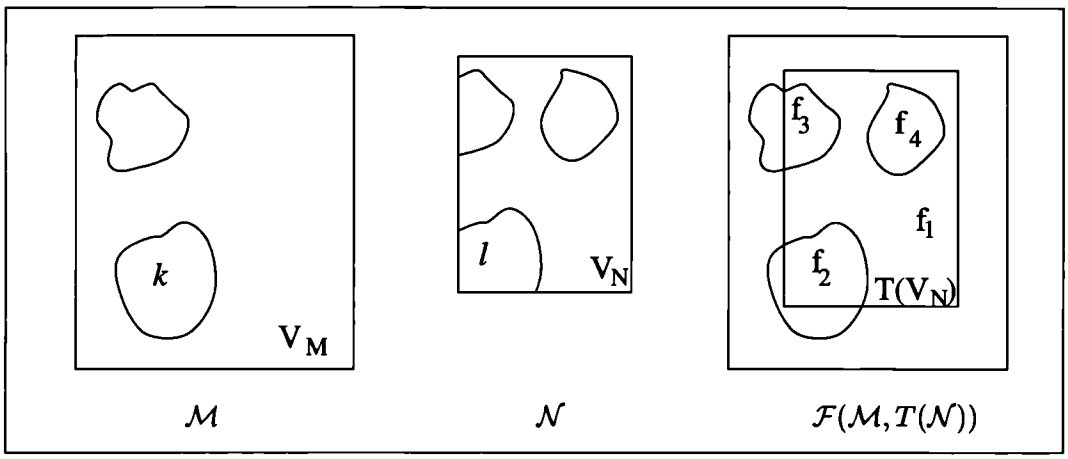


Figure 4.7: A sketch illustrating the truncation of imaged regions by the overlapping field of view.

cases, there will be regions for which the volume of overlap $\nu(f_2)$, where $f_2 = k \cap T(l)$, is constrained by the overlap of fields of view V_M and V_N of the two images.

As the transformation T varies the volume of overlap of regions k and l will vary both because of their alignment T, and because of their location within the overlapping field of view $(V_M \cap T(V_N))$ of the two images. The normalisation of a measure with respect to the statistics derived from the field of view is particularly important for 3D medical images where images with limited extent in one or more axis are common.

4.4.1 Image Overlap: A Simple Model

In order to look at the behaviour of the measures to the overlap problem, it is possible to construct a simple model describing a clinically plausible scene. In this case we examine identical views of the scene since we are simply interested in the affect of field of view and not differences due to modality.

Here we can describe the 2D alignment of images of a half circle (say a coronal cross section through the brain) imaged with different fields of view, as illustrated in Figure 4.8. Both images have an extent of 1 unit vertically. Horizontally, the first image, for

simplicity, has an infinite extent, while the other has a limited extent determined by the parameter F_{OV}. A half circle is placed at the centre of the rectangular field of view. We consider in this model the response of different measures to the change in in-plane rotational alignment θ, around the centre of the field of view (and circle).

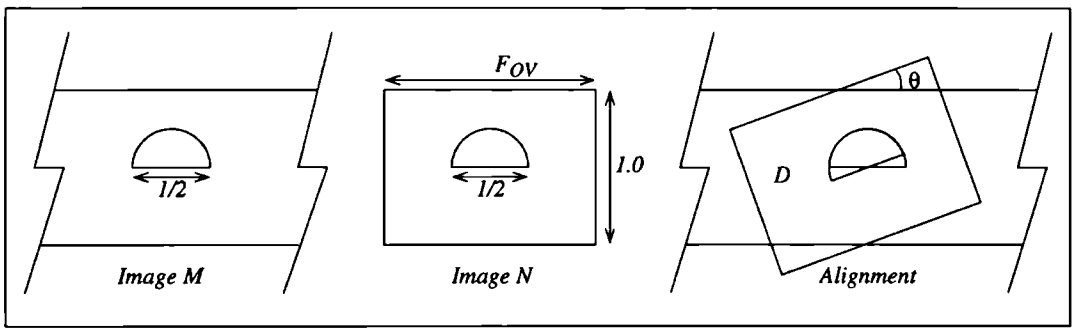


Figure 4.8: A simple model of rotational alignment θ between two images of a half circle with varying overlap and horizontal field of view determined by F_{OV}.

To evaluate the measures we need to calculate the area of overlapping regions in the two images as a function of the rotational alignment θ as shown in Figure 4.8(a). We can do this by considering the non-overlapping corners of the rectangle in Figure

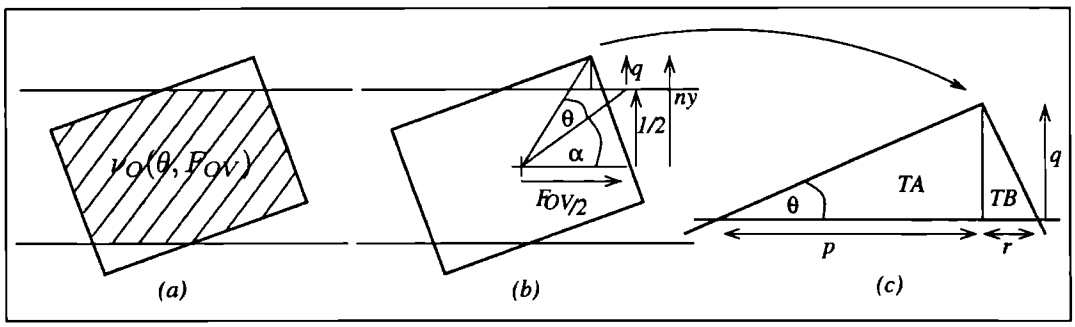


Figure 4.9: Evaluating the overlapping field of view of the two images for a given rotation θ and field of view F_{OV}.

4.9(b) made up of the two triangles TA and TB shown in 4.9(c). The area of these are determined by the change in the vertical position of the corner of the field of view q, given by $q = ny - \frac{1}{2}$,

$$q = \sin(\alpha + \theta)\sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{F_{OV}}{2}\right)^2} - \frac{1}{2}, \qquad 0 < \theta < 2\alpha \tag{4.14}$$

where $\alpha = tan^{-1}(\frac{1/2}{F_{OV}/2})$. For simplicity of the model, we can constrain the range of θ so that the other two corners of the rectangle (at the far left and right in figure 4.9(a))

remain inside the region of overlap. Given this we can then calculate p for the triangle TA using $\tan(\theta) = q/p$ which gives the area of TA as,

$$\nu_{TA}(\theta) = \frac{pq}{2} = \frac{q^2}{2\tan\theta} \tag{4.15}$$

Similarly the area of triangle TB can be shown to be,

$$\nu_{TB}(\theta) = \frac{rq}{2} = \frac{q^2\tan\theta}{2}. \tag{4.16}$$

From this, the overlapping area for rotation θ is given by $1.F_{OV} - 2(\nu(TA) + \nu(TB))$ which can be rewritten as,

$$\nu_O(\theta, F_{OV}) = 1.F_{OV} - q^2\left|\tan\theta + \frac{1}{\tan\theta}\right|. \tag{4.17}$$

In addition, we can work out the overlapping and non-overlapping regions of the half circle (radius $\frac{1}{4}$) shown in Figure 4.10 which are simply,

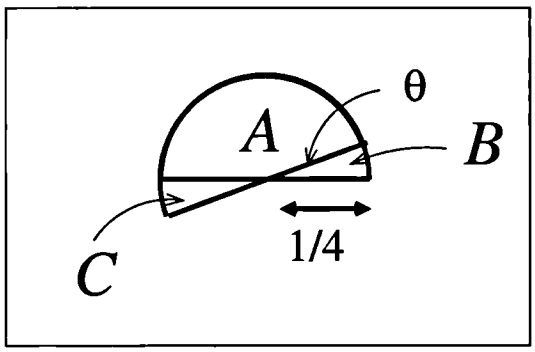


Figure 4.10: Evaluating the overlapping objects (half circles) in the two images for a given rotation θ.

$$\nu_A(\theta) = \frac{1}{2}\pi\left(\frac{1}{4}\right)^2\left(1.0 - \left|\frac{\theta}{\pi}\right|\right) = \frac{\pi - |\theta|}{32}, \qquad -\pi < \theta < \pi \tag{4.18}$$

$$\nu_B(\theta) = \frac{1}{2}\pi\left(\frac{1}{4}\right)^2\left|\frac{\theta}{\pi}\right| = \frac{|\theta|}{32}, \qquad -\pi < \theta < \pi \tag{4.19}$$

$$\nu_C(\theta) = \frac{1}{2}\pi\left(\frac{1}{4}\right)^2\left|\frac{\theta}{\pi}\right| = \frac{|\theta|}{32}, \qquad -\pi < \theta < \pi \tag{4.20}$$

The area of the 'background' region is then simply these subtracted from the overlap area,

$$\nu_D(\theta, F_{OV}) = \nu_O - \nu_A - \nu_B - \nu_C. \tag{4.21}$$

From these we can then evaluate the probabilities of each region (by dividing by the area of image overlap) and therefore the joint probability distribution, as a function of the transformation parameter θ and the field of view parameter F_{OV},

$$P(\theta, F_{OV}) = \frac{1}{\nu_O}\begin{bmatrix} \nu_A & \nu_C \\ \nu_B & \nu_D \end{bmatrix}. \tag{4.22}$$

Approaching alignment, $\theta \mapsto 0$ and

$$P(\theta, F_{OV}) \mapsto \frac{1}{F_{OV}} \begin{bmatrix} \frac{\pi}{32} & 0 \\ 0 & F_{OV} - \frac{\pi}{32} \end{bmatrix} \tag{4.23}$$

4.4.2 Response of Entropy Measures to Varying Field of View

Using this simple model of areas of overlapping regions in a pair of images, response of measures can be evaluated directly for different rotational misalignment and field of view. Plots for rotation θ between -30^o and 30^o, and field of view parameter F_{OV} between 0.5 and 6 are shown in Figures 4.11,4.12 and 4.13. As can be seen, the response of the joint entropy measure is affected considerably by the variation in field of view. As the field of view increases the value of the joint entropy increases and the peak at alignment is flattened. Going back to the axioms for information theory (section 4.2.1), minimisation of joint entropy favours images with a large field of view, if the background is already larger than the foreground. In this, case as the proportion of background increases, a value is more probably background than foreground and therefore the entropy falls.

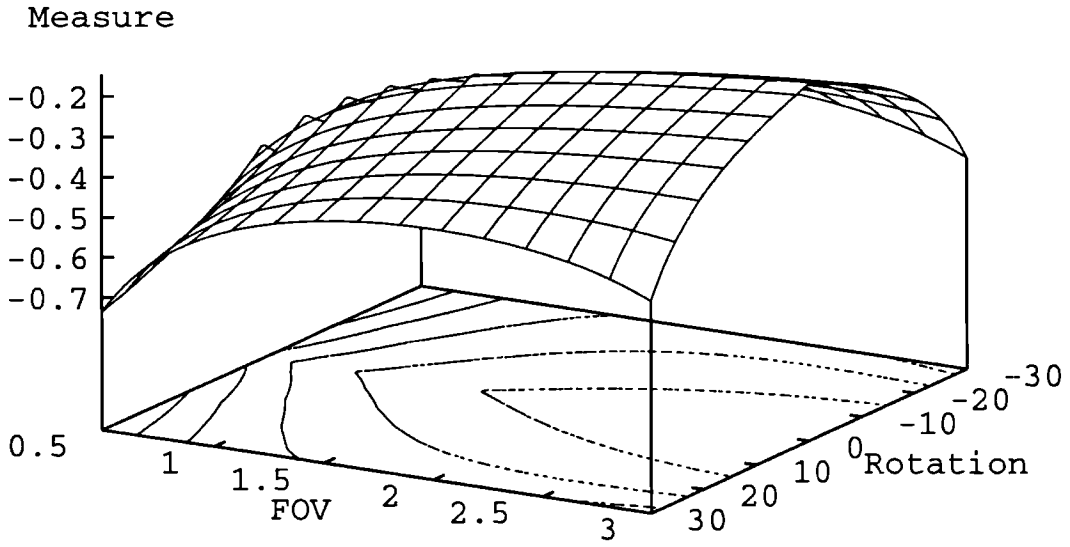


Figure 4.11: The response of negated Joint Entropy to rotational misalignment (degrees) at different values of field of view parameter F_{OV} of the model in Figure 4.8.

An even worse response with respect to field of view is exhibited by mutual information as shown in Figure 4.12. The response of mutual information is more complex than that of joint entropy, but can be simplified by considering the three components of $\frac{dI(M;N)}{dT}$ separately,

$$\frac{dI(\tilde{M};\tilde{N})}{dT} = \frac{dH(\tilde{M})}{dT} + \frac{dH(\tilde{N})}{dT} - \frac{dH(\tilde{M},\tilde{N})}{dT}. \tag{4.24}$$

As we move away from registration, joint entropy may increase so $\frac{dH(M\,N)}{dT} > 0$. But, as the overlap falls due to misalignment, $H(M)$ and $H(N)$ may not necessarily fall. In this case the 'background' region in the image is truncated as the images misalign. If this makes the background and foreground regions more similar in size then, as in Figure 4.1, the marginal entropies $H(\tilde{M})$ and $H(\tilde{N})$ will increase. So, if this increase is such that,

$$\frac{dH(\tilde{M})}{dT} + \frac{dH(\tilde{N})}{dT} \geq \frac{dH(\tilde{M},\tilde{N})}{dT}, \tag{4.25}$$

then mutual information does not provide a useful measure of alignment. In this model, the field of view increases to a point where this applies and the response to rotational alignment changes from a maximum to a minimum.

It is important to note that this condition applies to any image pair where the background and foreground regions and overlapping field of view are such that the marginal entropies are sufficiently affected by overlap.

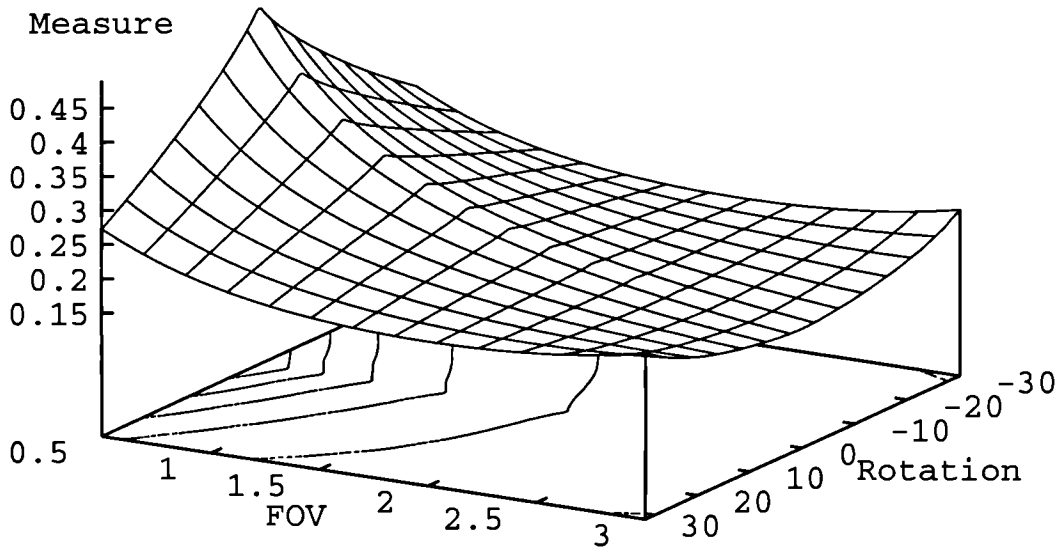


Figure 4.12: The response of Mutual Information $I(\tilde{M};\tilde{N})$ to rotational misalignment (degrees) at different values of field of view parameter F_{OV} of the model in Figure 4.8.

The normalised mutual information measure (Figure 4.13) provides a response virtually independent of the field of view parameter F_{OV}, with a much sharper exponential relationship between field of view and measure. Because here we are dividing the marginal entropies by the joint entropies, any increase in the marginal entropies will be counteracted by the change in the joint entropy.

As a comparison the two measures proposed by Maes *et al* [64] were also evaluated for this model. The entropy correlation coefficient, as would be expected from its mathematical form, provides a similar response to that of normalised mutual information

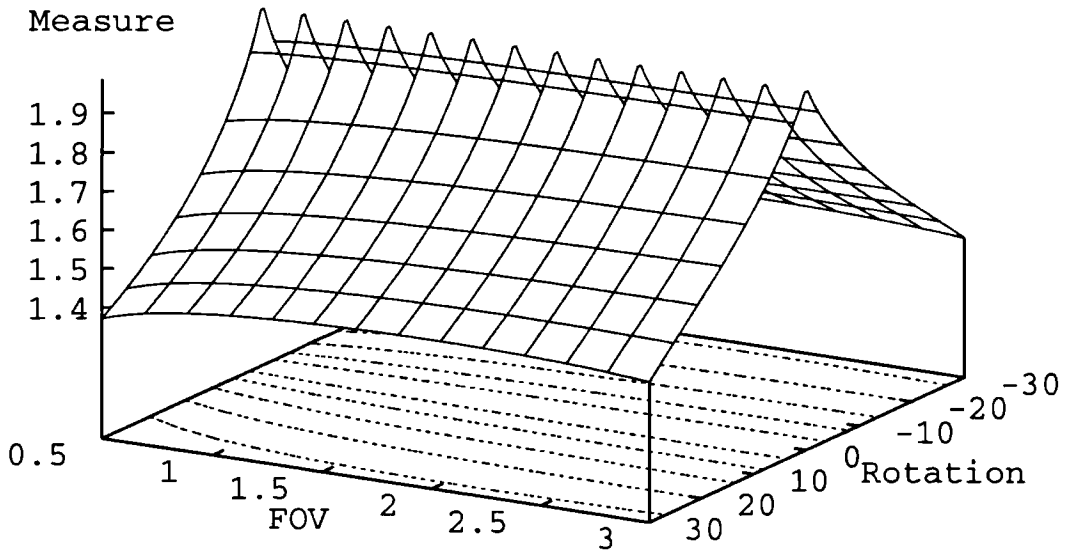


Figure 4.13: The response of Normalised Mutual Information $Y(\tilde{M};\tilde{N})$ to rotational misalignment (degrees) at different values of field of view parameter F_{OV} of the model in Figure 4.8.

(Figure 4.14). The response of $\rho(\tilde{M};\tilde{N})$ though, shown in Figure 4.15, is less sharp than that for either normalised mutual information or entropy correlation coefficient.

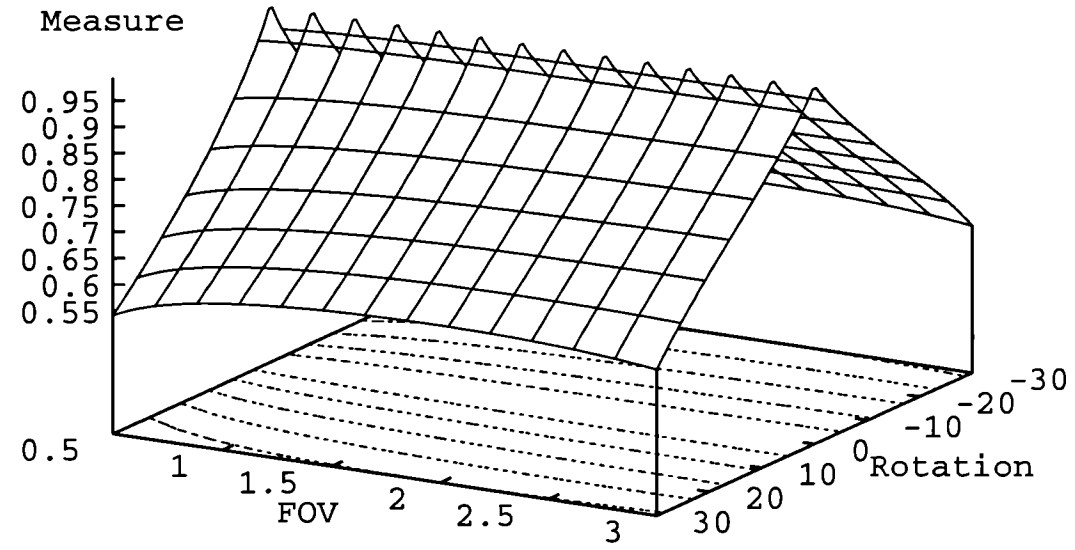


Figure 4.14: The response of the entropy Correlation Coefficient $C_e(\tilde{M};\tilde{N})$ to rotational misalignment (degrees) at different values of field of view parameter F_{OV} of the model in Figure 4.8.

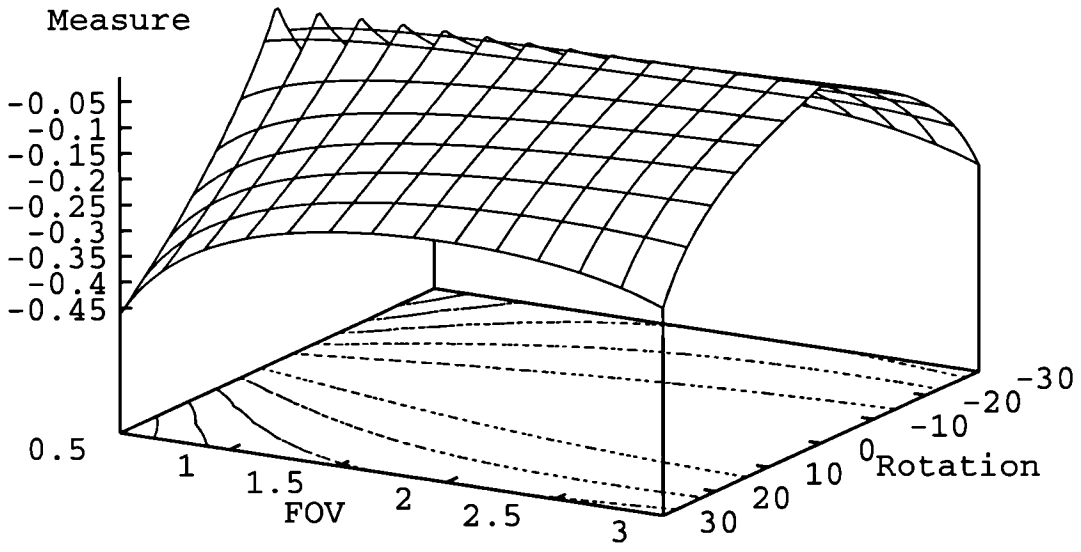


Figure 4.15: The response of the entropy measure $\rho(\tilde{M};\tilde{N})$ to rotational misalignment (degrees) at different values of field of view parameter F_{OV} of the model in Figure 4.8.

4.5 Summary

This chapter has described and developed the use of entropy as a generic measure of multi-modality alignment. By treating the images as sources of information, information theory provides the basic tools for describing and developing generic measures of alignment. Such an approach infers 'similarity' of image values only from the observed occurrence of those values together in the image as a whole, allowing us to relate and manipulate values of completely unrelated physical properties.

In using information theory to examine the spatial relationship between images, the problem of normalisation with respect to image content was identified as being an important factor. This is particularly true for our application where we want to compare alignment estimates between volumetric images with limited axial extent. By constructing a simple model of a scene, the response of the different measures as the field of view is varied, can be examined and compared. The normalised mutual information measure $Y(\tilde{M};\tilde{N})$ was identified as providing a robust measure of alignment, independent of the local image statistics. The results of this simulation will be shown to be applicable to clinical images in the experiments of chapter 6.

Chapter 5

Evaluation and Optimisation of Image Similarity

5.1 Introduction

The previous two chapters dealt with the ideas behind the development of similarity measures between views of a scene provided by different modalities, containing both corresponding and complementary structure. This chapter deals with the problems of implementation and use of these measures in recovering the alignment of *images* of those views. Rather than concentrate on the use of one particular measure this chapter is concerned with the generic problems of:

- Evaluation of corresponding measurement values.
- Evaluation of a joint probability distribution $p\{m, n\}$ of those values.
- Optimisation of an alignment measure between finite 3D images.

The chapter begins with a brief description of the effect of image noise and limited spatial resolution of the structure of the joint probability distribution. The next section then goes on to look in more detail at the problem of evaluating spatially corresponding measurements between 3D medical modalities. Building on this, the next section examines the problem of forming an estimate of the joint probability distribution for a given transformation estimate.

Having looked at these basic tools of registration evaluation, the chapter then goes on to look at the problem of finding the optimum alignment for a given measure, beginning with an examination of the factors affecting the structure of the alignment parameter space. A simple multi-resolution optimisation scheme is then described which makes use of efficient low resolution evaluation of registration measures.

5.2 From Properties to Measurements: Noise and Spatial Resolution

Up until now we have considered only an idealized description of the views of a scene provided by different modalities. The approaches to quantifying alignment are directly applicable to real clinical image data consisting of measurements limited by noise and spatial resolution.

The relationship between the image function $m(\mathbf{x})$ and the scene or object function $\tilde{m}(\mathbf{x})$ is governed by the point spread function $h(\mathbf{x})$ and noise function $\epsilon(.)$ of the imaging system. The relationship can generally be modeled [100] by:

$$m(\mathbf{x}) = \int h(\mathbf{x} - \mathbf{u})\tilde{m}(\mathbf{x})d\mathbf{u} + \epsilon_m(\mathbf{x}, \tilde{m}(\mathbf{x})). \tag{5.1}$$

5.2.1 Measurement Noise

In many cases measurement noise, can be approximated as being independent of spatial location[1] $\mathbf{x}$ so,

$$\epsilon_m(\mathbf{x}, \tilde{m}(\mathbf{x})) \simeq \epsilon_m(\tilde{m}(\mathbf{x})) \tag{5.2}$$

This uncertainty in the measured value m given the underlying value of the material $\tilde{m}$ can be modeled with a probability distribution, say,

$$p(m|\tilde{m}) = E_m(\tilde{m}, m). \tag{5.3}$$

The following sections briefly describe the form of the noise distribution in the main 3D imaging modalities investigated in this thesis.

Noise in CT images

In CT, the image formed represents the linear X-Ray attenuation coefficient μ of material. The detector noise in a CT scan is due to quantum fluctuations in the X-Ray beam and noise in the detector [14] and has a Poisson distribution. In this case the probability of recording a measurement of value m (in this case Hounsfield Units) given that the imaged material has an actual value $\tilde{m}$ is given by,

$$p(m|\tilde{m}) = \frac{\lambda^{|m-\tilde{m}|}}{|m - \tilde{m}|}\exp(-\lambda). \tag{5.4}$$

In practice, because λ is large this can be approximated as a simple Gaussian distribution,

$$p(m|\tilde{m}) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(\tilde{m} - m)^2}{2\sigma^2}\right\}. \tag{5.5}$$

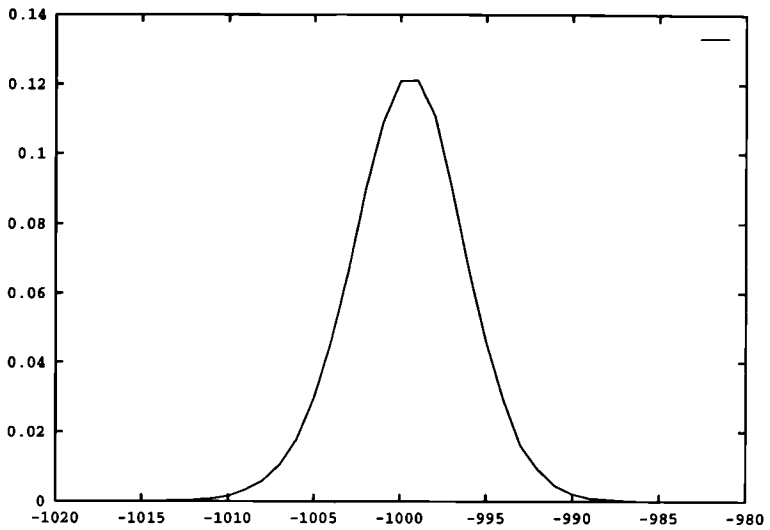


Figure 5.1: A distribution of CT image values from a uniform region of material (air value -1000, data stored in Hounsfield units)

A distribution derived from an area in one of the images used in the later experiments (patient 5 of the Vanderbilt database) is shown in figure 5.1. This distribution varies depending on the resolution of the image, but the main source of additional artefact in the distribution is the truncation of the dynamic range due to its storage in binary form. It is common for data to be stored in 'CT numbers' which are effectively Hounsfield numbers with 1024 added to make -1000 Hounsfield units (air) equal to 24 CT 'units'. This, depending on the spread of the noise distribution, does not remove all negative values from the image. As a result any CT numbers less than zero are set to zero. A side affect of this is the apparent noise level in the air is reduced, reducing particularly the visibility of reconstruction artefacts in the regions of air around the patient. An example distribution where this occurs is shown from a region of air in a CT image of the pelvis acquired on a Siemens CT scanner in Figure 5.2.

The portion of truncated values varies with the spread of the distribution (and therefore the resolution of the images). Other less significant affects are caused by alternative methods of image reconstruction, in particular spiral CT reconstruction. An example distribution from such an image is shown on the right of figure 5.2.

Noise in MR Images

In MR, the property effectively being measured is the magnetic moment density. The image is formed by a Fourier reconstruction process from complex data. The noise in the complex signal data can again be modeled as simple uncorrelated Gaussian noise added

[1]Later chapters will address a case where this assumption does not hold

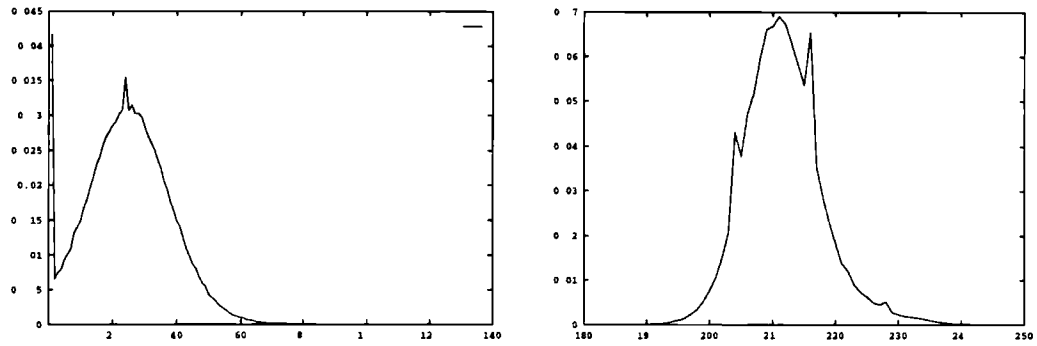


Figure 5.2: Artefactual CT Noise Distributions from a region of air: Left (value 24) showing truncation of the dynamic range below zero due to storage. Right: Artefacts in noise distribution from a region of air in a Spiral CT acquisition, possibly due to the method of reconstruction.

to the real and imaginary components. The actual clinical MR image is a modulus of this complex Fourier transformed signal. The resulting distribution of additive noise in the modulus image is dependent on the underlying signal level [62]. For zero signal (e.g. in air) the noise in the modulus image can be shown to have a *Rayleigh* distribution. At higher underlying signal levels, the distribution is *Rician*, tending to a Gaussian distribution. For many tissues a Gaussian distribution is an adequate model. The distributions of noise in a modulus image of air and a uniform solution of $CuSO_4$ are illustrated in figure 5.3.

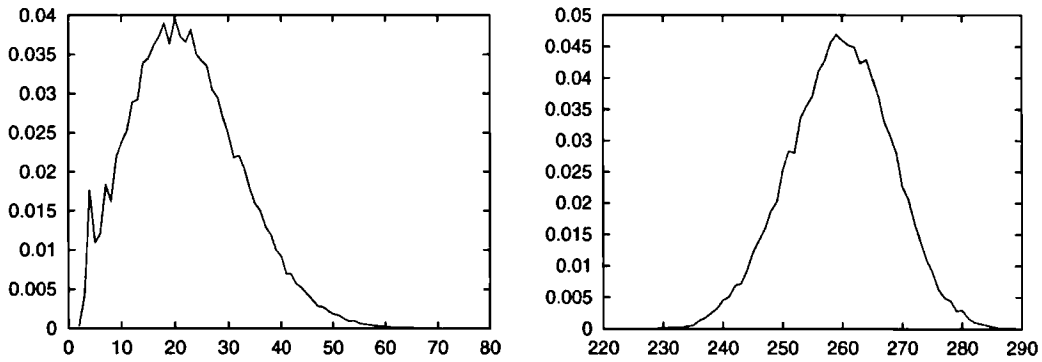


Figure 5.3: Noise distribution from a region of air (left) and and region of uniform solution (right) in a modulus MR image.

Noise in PET Images

In PET imaging the measurements making up an image effectively represent the radio-isotope density due to tracer uptake at a given location. The inherent detected noise, as with X-Ray CT has an approximately Gaussian distribution. A distribution of values from a region of air in a PET image is illustrated in figure 5.4. This also shows another

example of corruption of the data in the reconstruction or storage process of the scanner software. Here there is a peak roughly twice the expected height at zero. This is most probably a result of the use of incorrect rounding (e.g. 'C' int() rather than nint()) when the scanner software converts the reconstructed floating point data to integer format for storage.

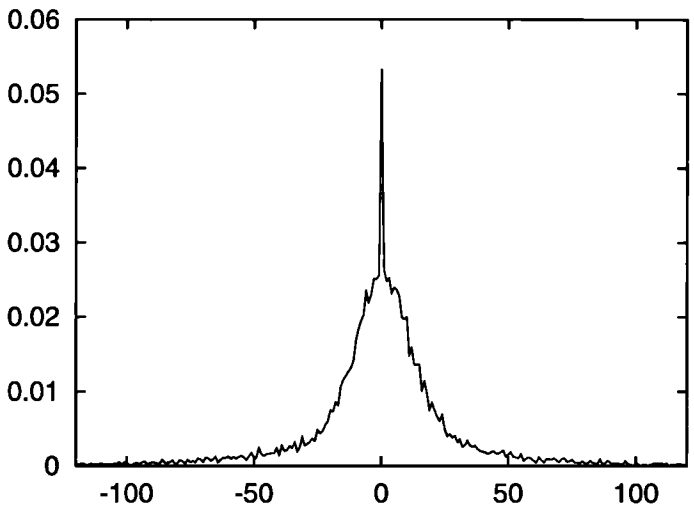


Figure 5.4: A distribution of image values from a uniform region of air in a PET scan illustrating the noise distribution. Increased peak at zero may be due possibly to incorrect rounding when converting the image data to integer numbers in the reconstruction software.

Quantization Noise

A secondary source of noise is that of quantization due to the discretization of measurement values. This is determined by the gain controls and the number of bits of the analogue to digital converter (ADC) of the imaging system. A common design approach in medical imaging systems is to employ an ADC with enough resolution (number of bits) to ensure that the least significant bit is simply representing the noise level in the image measurements.

Noise in the Joint Probability Distribution

Given the noise in the two images will be independent, for a given transformation, the joint probability distribution of corresponding *measured* values $p(m,n)$ in the two images is simply the distribution of actual material properties $p(\tilde{m},\tilde{n})$ convolved with the noise distributions for each modality,

$$p(m,n) = \int\int p(\tilde{m},\tilde{n}) E_m(\tilde{m},m) E_n(\tilde{n},n) d\tilde{n} d\tilde{m}. \tag{5.6}$$

Examples of distributions of corresponding MR-CT and MR-PET values from brain images are shown in figures 5.5 and 5.6. These are displayed as a surface whose height represents the percentage of the volume exhibiting a given pair of values in the two images. The surface height has been truncated to exclude the peak due to the most probable values corresponding to air (the lowest value in both modalities) to illustrate the distribution due to smaller structures. Note particularly the considerably increased spread of the PET values due to the higher level of noise in this modality.

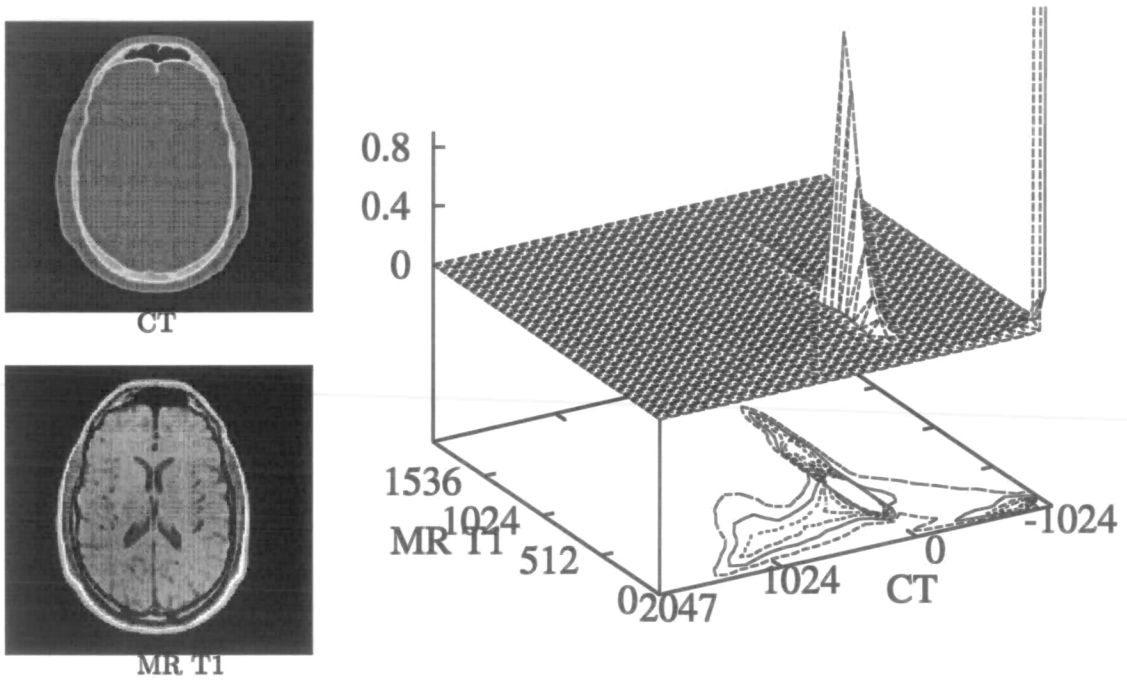


Figure 5.5: Example of a joint probability distribution of values in a T1 weighted MR image (MR modulus values) and a CT (Hounsfield units) of the brain (alignment derived from bone implanted markers).

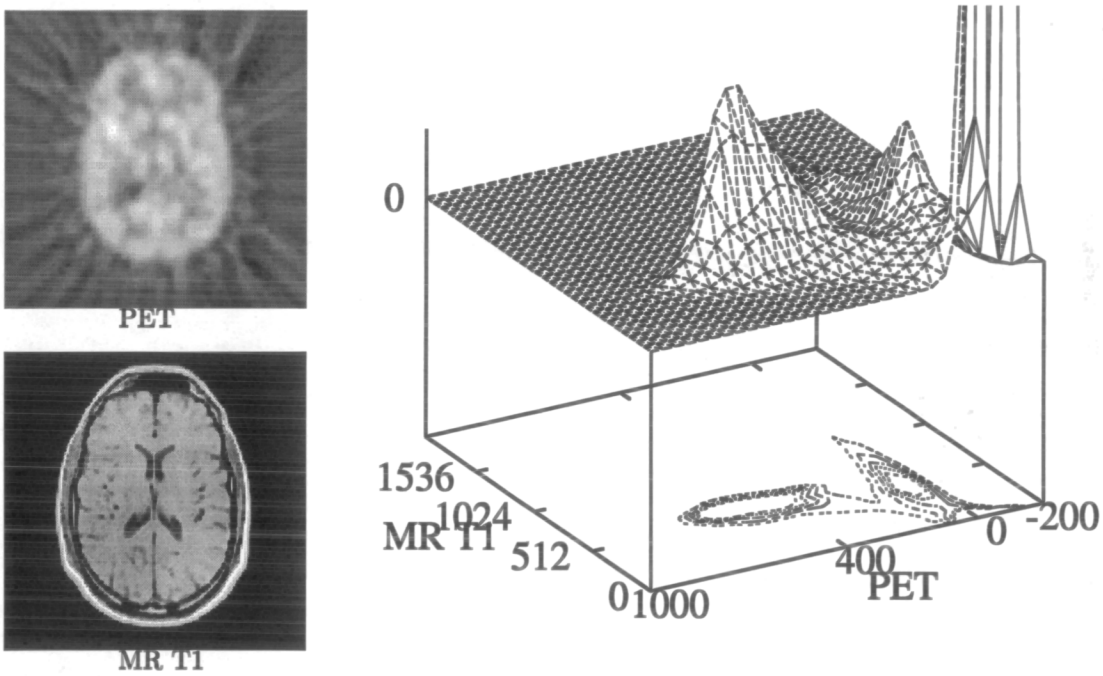


Figure 5.6: Example of a joint probability distribution of values in a T1 weighted MR image (MR modulus values) and a PET FDG image of the brain (alignment derived from bone implanted markers).

5.2.2 Measurement Resolution

In real imaging systems, finite *spatial resolution* due to the impulse response function $h()$ in equation (5.1) limits our knowledge about the anatomical and physiological regions within the patient. A measurement in an image is therefore derived from some combination of properties of a region of material in the scene. In general, because of the different physical processes used in imaging, the point spread function $h()$ can vary considerably between different modalities.

The images are made up of measurements at discrete locations $\mathbf{x} \in V_M$ (the *sampling resolution*) which are governed by the imaging system and the image reconstruction parameters. The fact that the impulse response $h()$ has an extent greater than the sampling interval means that neighbouring image measurements $m(\mathbf{x})$ and $m(\mathbf{x} + \delta\mathbf{x})$ contain contributions from the same locations and therefore are spatially correlated. At boundaries between different materials, a measurement may have contributions from the properties of both materials, resulting in the so called *Partial Volume* effect. A measurement value for example, in a CT image, corresponding to soft tissue may be recorded at a boundary between regions of bone and air. The overall effect of the point spread function on the joint probability distribution $p\{m, n\}$ is complex and dependent on the structures within the scene being imaged.

One of the main problems in multi-modality imaging is that it is common for the two images to have significantly different resolutions. If one image has a higher spatial resolution than the other, then small objects which are delineated in the higher resolution image may not be delineated by the lower resolution image. A region which directly corresponds in the two modalities when imaged at the same resolution, may exhibit a much more complex boundary at a higher resolution. A useful way of looking at this affect is in terms of the iso-intensity contours in the two images as the resolution changes.

As the spatial resolution of one modality is reduced, the added structure at region boundaries means that it is no longer possible to define a threshold which delineates the same contour in both modalities as illustrated in figure 5.7. In terms of the joint probability distribution, a single peak representing the object at equivalent resolutions will disperse to form multiple peaks as one image is blurred.

Resolution and Alignment

As we vary the transformation estimate and evaluate a measure, each single point correspondence contributes a 'count' to the joint probability distribution. The change in measurement values $\frac{dm(\mathbf{x})}{d\mathbf{x}}$ and $\frac{dn(\mathbf{y})}{d\mathbf{y}}$ across the images determines the change in location of each contribution to the joint probability distribution as the transformation T varies.

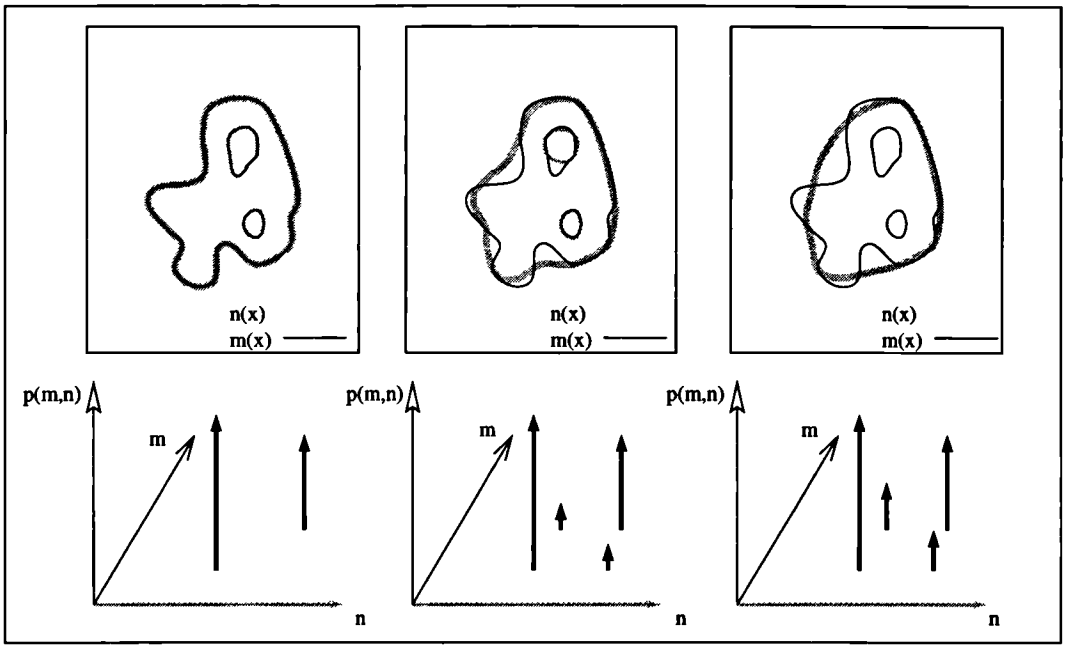


Figure 5.7: A sketch illustrating the effect of blurring one image $n(\mathbf{x})$ with respect to a second $m(\mathbf{x})$ on the joint probability distribution of corresponding values $p\{m,n\}$. Images are represented by iso-contours of corresponding values at three levels of blurring. Blurring introduces peaks into the joint probability distribution where there are non-corresponding regions due to changes in the location of a boundary.

We can examine this in the simple terms of the mapping of a single value $m(\mathbf{x})$ in one image to a location in the other image as shown in Figure 5.8. As we reduce the spatial resolution of the images, so we reduce the rate of change of the probability distribution,

$$\frac{p\{m,n\}}{dT} = \begin{bmatrix} \frac{p\{m_1,n_1\}}{dT} & \cdot & \cdot & \frac{p\{m_1,n_j\}}{dT} \\ \cdot & \cdot & \cdot & \cdot \\ \frac{p\{m_i,n_1\}}{dT} & \cdot & \cdot & \frac{p\{m_i,n_j\}}{dT} \end{bmatrix} \tag{5.7}$$

with respect to changes in the transformation. This then reduces the rate of change of any measure of alignment derived from this distribution.

5.3 Evaluating Spatially Corresponding Measurements

The most fundamental task in evaluating the alignment of discrete images is forming an estimate of measurements at corresponding points for a given transformation estimate. Different imaging systems may derive an estimate of the value of a property from different volumes of material as illustrated in figure 5.9 (top). Ideally we would like to relate measurements of different properties, not only at a corresponding location, but also over an equivalent volume of material at that location.

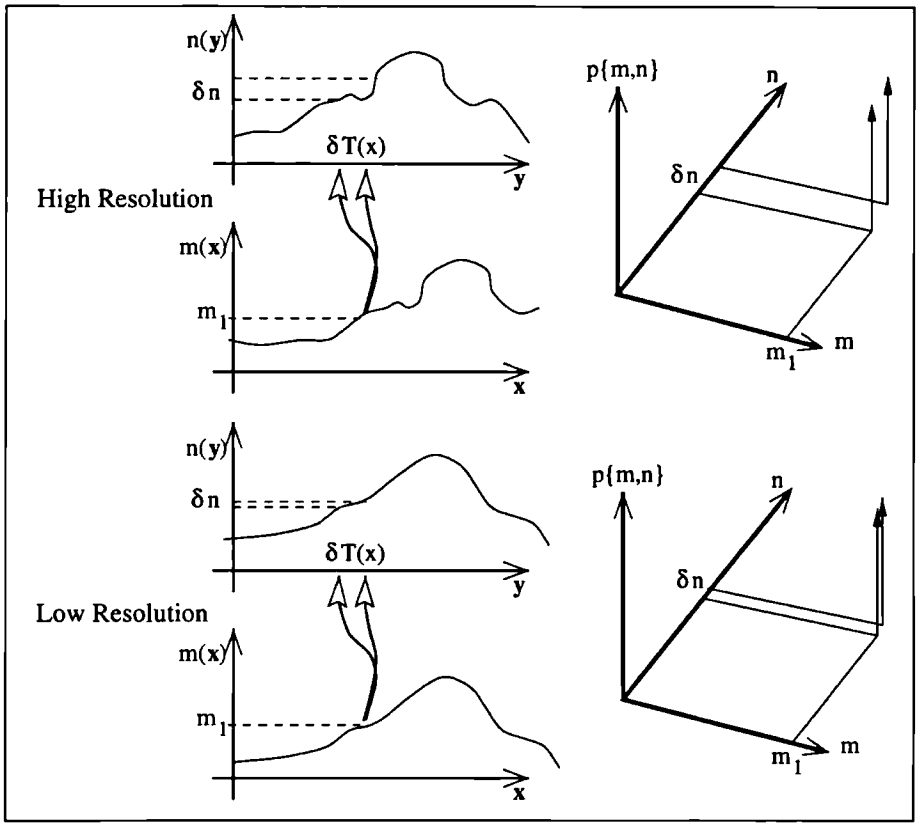


Figure 5.8: Sketch illustrating the effect of lower spatial resolution on the change in contributions to the joint probability distribution $p\{m,n\}$ between images $m(\mathbf{x})$ and $n(\mathbf{y})$, as the transformation estimate $T : x \mapsto y$ is varied by an amount dT. As the resolution is reduced the displacement of a point $dT(\mathbf{x})$ results in a smaller change in corresponding intensity $n(T(\mathbf{x}))$.

In practice we have a continuous rigid body transformation between two discretely sampled images. As a result, we also need to consider the problem of finding, for a given measurement in one image, a corresponding value at a location which does not fall on a sampling point of the second image. This is a problem of interpolation.

5.3.1 An Engineering Compromise: Two Stage Resampling

Resampling so that a truly equivalent measurement is estimated at a corresponding location in the two images is both mathematically complex and computationally expensive. In order to carry out a registration we need to iteratively optimise measures. This requires many evaluations of the joint probability distribution. In practice, the large volume of data means that an estimate of the probability distribution from two images will require many evaluations of corresponding values [2]. Given the practical limitations

[2] there may be 8 million measurements in a typical 3D MR volume acquisition of the brain

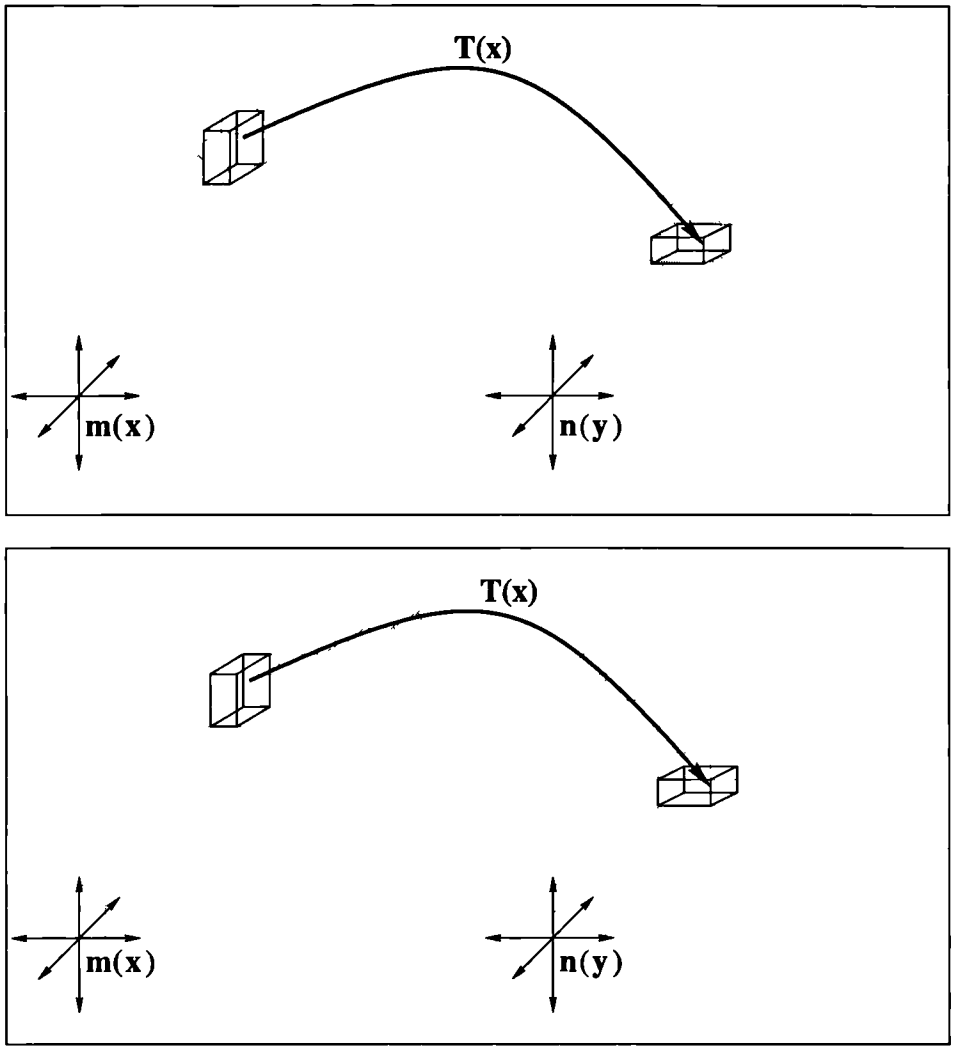


Figure 5.9: To relate values in the two modalities we need to take into account the points at which measurements are made (the sampling rate, bottom) and the volume of material over which the different measurements originate (sampling resolution, top)

in computing resources and a final aim for a system to run on mid-range workstations, we want to keep the number of operations required to evaluate corresponding values to a minimum.

A key factor in evaluating corresponding measurements is the cost of carrying out either a bandlimited reduction in sample rate or accurate interpolation. The approach taken in this work is to divide up the sampling problem into two stages. A resampling step prior to alignment ensures the images have comparable spatial resolution and sampling rate in each axis (Figure 5.10). Inexpensive linear interpolation is then used to evaluate the joint probability distribution during iterative registration.

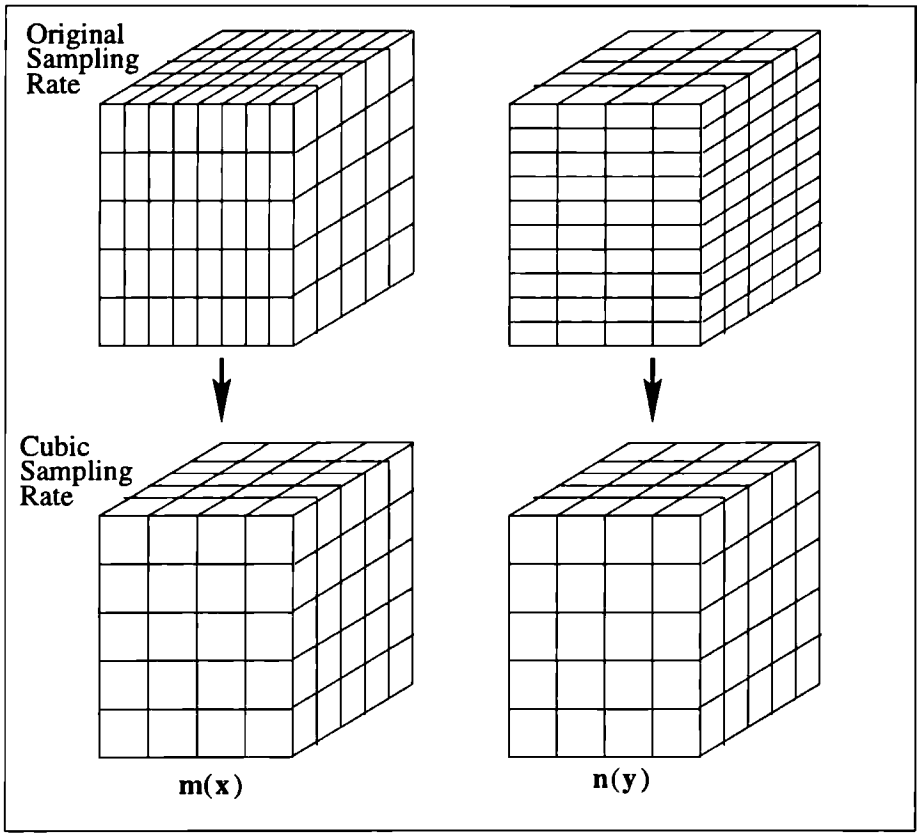


Figure 5.10: An illustration of cubic resampling of 3D images prior to alignment.

5.3.2 Isotropic Resolution

In addition to isotropic sampling we would like to impose a similar, isotropic spatial resolution on the measurements in the two images. In practice we cannot increase the spatial resolution of an image. We can though reduce the spatial resolution of images where one is appreciably higher resolution than the other. The main case where this occurs is where we are aligning for example MR images with PET or SPECT images. Here the PET or SPECT can have in impulse response $h()$ with a full width half maximum (F.W.H.M) of 15mm or better, where a modern 3D MR of the brain can have a FWHM of around 1mm.

The approach taken here is to apply a low pass filter kernel to the higher resolution image to increase the width of its impulse response so that it is comparable to that of the lower resolution image. Many of the responses of the lower resolution modalities such as PET can be approximated by a Gaussian [5] and so in these experiments on MR-PET registration a Gaussian filter kernel,

$$G(\mathbf{k}, \sigma) = \frac{1}{(\sigma\sqrt{2\pi})^d} \exp\left(\frac{-\mathbf{k}^2}{2\sigma^2}\right) \tag{5.8}$$

was used to filter the MR images prior to alignment. Here $\mathbf{k}$ is the distance in sample units in the source image from the interpolating point, and d is the dimensionality which in our case is equal to 3.

Field of View and Data Padding

For many modalities such as CT and some nuclear medicine imagery, the reconstructed field within a plane through the patient is circular. This results in an overall cylindrical image volume which is commonly stored in a rectangular array by padding the border with a constant value. At best, for images where the section of the patient is surrounded by air in the images, there will be a synthetic edge in the 'rectangular image' at the boundary of the air and padding region. For clinical images reconstructed with a smaller field of view, there will often be a direct truncation of the patient tissues by the circular reconstruction field as illustrated by the clinical example in figure 7.6 in chapter 7.

Regions of padding can have two effects on the joint probability distribution and therefore the similarity measures. Firstly the uniform region of padding does not possess the same statistical noise properties as the actual image data, resulting in a discontinuity in the joint probability distribution. Secondly we must be careful when subsampling and interpolating the images not to allow any padding values to corrupt the true image measurements at the image-padding boundaries. This is particularly important where there are soft tissue or bone regions at the image border.

In some image formats, the padding regions are given a value distinct from actual measurement regions, and so can be easily distinguished. Commonly though this is not the case and padding regions are labelled with a value 0 which is equivalent to the measurement exhibited by air in many CT and MR datasets. As a result, there is no easy way to distinguish measurements from padding. The approach taken in this work is to extract a padding region as a uniform connected set of voxels around the edges of the rectangular volume. A padding value is simply detected by the presence of the same value along an edge of the rectangular volume between two of its vertices. If this occurs then it is most probably padding (since if it were real signal there would be noise on the measurements and there would not be a continuous value along the edge). A connected component fill is then carried out to label all the padding voxels in from this edge with a value distinct from the actual image data values. These values are then ignored in any further subsampling or evaluation.

Finite Images and Kernels

For any image resampling process there is the problem of handling the finite extent of the data. In kernel based resampling this means estimating values at points where the edge of a kernel extends beyond the boundary of the recorded data. This is a common problem when manipulating rectangular image data, but for many tomographic modalities the problem is even more complex.

There are many proposed approaches to handling the boundary problem [98] with varying levels of complexity. For this work a simple truncation of the subsampling kernel area was used which is simple to apply to the cylindrical image volumes. The kernel summation is evaluated at all valid points in the source image. If the area of the kernel falling on a valid image measurement exceeds 50% of the total kernel area then the value is normalised for the fraction of kernel used, and taken as a valid subsampled value in the destination image. Otherwise the subsampled point is set to the 'padding' value in the subsampled image and ignored in further computations. The affect of using different overlap thresholds for a Gaussian kernel when subsampling an image of stripes with a circular field of view are shown in figure 5.11.

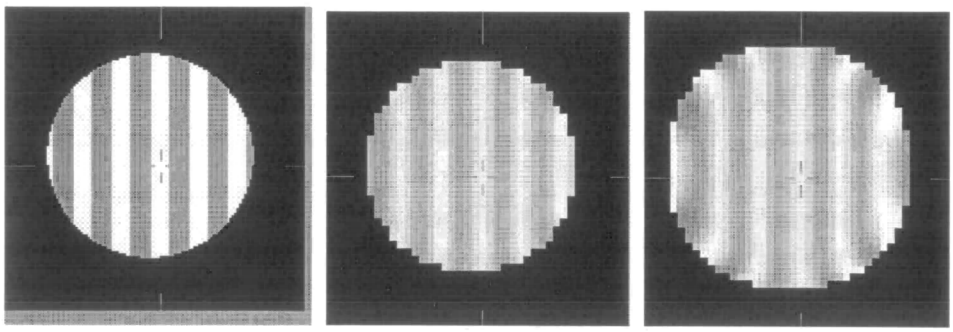


Figure 5.11: The effect of varying kernel overlap threshold (middle:50% and right:0%) on subsampling of an image of stripes (left) with a circular field of view.

5.3.3 Interpolation During Iterative Alignment

Having resampled both the modalities to cubic voxels of the same dimension, we must then evaluate corresponding values for a continuous transformation estimate T between these two image volumes. Any approach to this interpolation problem must supply a solution that requires minimal processing since it is this stage which governs the computational expense of alignment.

The accuracy of an interpolated value is governed by the extent of the kernel support used to derive the estimate. The most accurate interpolation is achieved with a sinc [12]

kernel of large extent. As the kernel size is reduced the interpolation accuracy falls and alternative interpolation functions become competitive. These include the use of cubic spline based kernels and the Keys cubic convolution kernel [92]. Even these kernels in three dimensions require considerable numbers of multiplication-addition operations. A 4^3 kernel evaluated over a 100^3 voxel image requires more than 64 million such operations (neglecting loop overheads). An iterative search scheme may require many hundreds of evaluations to locate an acceptable optimum. As a result it was decided to limit interpolation overhead by using only 2^3 kernel (8 million operations). Given the limited size, the choice of kernel was limited and in this case a simple piecewise linear kernel was employed [92]. Considering only the one dimensional case, if a value is to be estimated at a point x, then we can evaluate contributions of neighbouring sample points at a distance x_k from this point using

$$\beta^1(x_k) = \begin{cases} 1 - |x_k|, & |x_k| < 1 \\ 0, & \text{otherwise} \end{cases} \tag{5.9}$$

Where the distance x_k is in fractions of sample units. The kernel is separable and in three dimensions the weighting at a distance $\mathbf{x}_k = \{x_1, x_2, x_3\}$ is given by,

$$\beta_3^1(\mathbf{x}_k) = \beta^1(x_1)\beta^1(x_2)\beta^1(x_3). \tag{5.10}$$

One of the main limitations of this function is that the intensity gradient of the interpolated image is not continuously varying between sample points. In addition, the kernel also tends to act as a low pass filter, smoothing out any high frequency structure. When the sampling grids are aligned, the kernel provides an exact copy of the input image.

5.4 Forming an Estimate of the Joint Probability Distribution

Given we can evaluate corresponding values in an image, the next step is to form an estimate of the joint probability distribution of these values $p\{m, n\}$. Essentially there are two approaches to this problem. We can either form a continuous mathematical estimate of the distribution by, for example, fitting a function to the data. Alternatively, we can form a discrete histogram of the distribution from the data. Both approaches have been proposed and investigated for multi-modality image alignment.

5.4.1 Discrete versus Continuous

The most common approach to forming a continuous estimate of the distribution is using a Parzen Window approach [28]. This has been proposed by Collignon *et al.* [21] and

Viola *et al.* [99] as an approach to evaluating registration. Collignon argued that the computational expense of the Parzen window estimate means that for medical data it is much easier to simply form a discrete estimate of $p\{m,n\}$. Viola [99] on the other hand described and implemented an elegant Parzen window estimation in conjunction with a stochastic sampling and optimisation scheme for multi-modal registration. This essentially models the distribution as a combination of Gaussians, using an entropy driven maximum likelihood estimate. Computational efficiency was achieved through the use of significant sub-sampling of the data to form an estimate from a small number (as small as 50) of corresponding data points. In later work this was extended using a multi-resolution scheme to provide a more robust registration [102].

The published results using the stochastic sub-sampling scheme indicated that the final precision of estimates were limited [102]. Higher precision may well require substantially more points to be used in the Parzen estimate and therefore require substantially more computation. This computational limitation suggests an important question: What do we require of our estimate of $p\{m,n\}$? A computationally accurate and continuous estimate of the distribution may not be needed simply to assess image alignment. In general there will be a relatively small number of distinct materials and therefore properties delineated by a modality (for example grey matter, white matter, CSF, bone, air, fat). These regions of material will in general be uniform and, at imaging resolutions, have distinct boundaries (although the transitions may be blurred at, for example the cellular level).

The alternative approach of forming a discrete histogram of image values which is used in this thesis, was proposed earlier by Hill *et al.* [46, 48] and Collignon *et al.* [21]. This simply makes use of a direct count of the number of particular pairs of binned values occurring together. Since no functional fit is required to the data it is computationally less expensive when derived from a large number of data points.

In the scheme employed in this thesis, the histogram is derived by fully sampling points in the images at a given spatial resolution. As described later, computational efficiency is achieved, not through random or regular subsampling of images, but through fully sampling lower resolution (bandlimited) versions of the images.

5.4.2 Binning Values

Given we want to form a discrete 2D histogram of the joint occurrence of image values, the question is which image values do we use ? If there was no noise, then we would only need to consider the occurrence of values of the finite number of materials in a scene. In practice, we do not know exactly what materials will be present and for most cases

exactly what values they will exhibit. CT images are well calibrated with particular values corresponding to particular tissues, but the values in MR and PET images are much less standardised. We can though simply consider a typical range of values as illustrated in Figure 5.12. We must though decide how we wish to partition this range $\hat{M} = \wp(M)$.

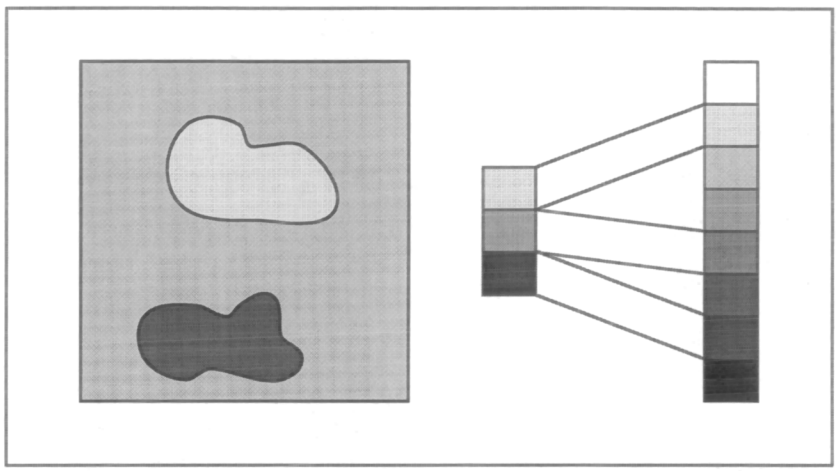


Figure 5.12: Because of unknown anatomy, physiology and limited calibration of some modalities, we cannot assume the specific properties of materials present in the images, so we consider a range of values.

Empirically we should use enough bins so that distinct materials delineated by values in the imaging modalities contribute to different bins in the histogram. For registration purposes we need in practice only to delineate region boundaries which are useful in registration (i.e. that appear in both images). This therefore imposes a lower limit on the number of bins we should use. An upper limit is imposed by the discrete number of levels, derived from the number of bits of the ADC, used to record the data.

For the experiments in this thesis, unless otherwise stated a total of 64 bins over the full range of image intensities was used for both modalities. Experiments in later chapters verify that varying the number of bins between 32 and 256 has little affect on the accuracy or precision of the final registration estimate provided similarity measures.

5.5 Alignment Search Space

Given a transformation T mapping points in an image $m(\mathbf{x})$ of volume V_M, to points in an image $n(\mathbf{y})$ of volume V_N, we need to find the value of T_r,

$$T_r = \max_T \{\mathcal{S}(m, n, T)\}, \tag{5.11}$$

such that the similarity $\mathcal{S}$ is optimised (a minimum or maximum depending on the measure).

One of the major differences in the optimisation problem posed by corresponding object alignment and similarity measure based approaches, is that any measure of image similarity can only be derived from a finite region of overlap,

$$V_{MN}(T) = V_M \cap T(V_N) \tag{5.12}$$

of the two imaged volumes. As misregistration continues to increase, the volume of overlap may decrease. The quantity of image structure over which the registration measure is evaluated therefore falls. In solving the large scale alignment problem we may compare measures derived from very different volumes of overlap. The data within the overlap at registration T_r may be unrelated to that at alignment, i.e.

$$V_{MN}(T_r) \cap V_{MN}(T) \mapsto \emptyset. \tag{5.13}$$

At extreme misregistration, depending on the content of the images, a measure may be better than that at true registration as illustrated in figure 5.13. Thus, in searching for true registration we may really only be seeking to find a local optima of any measure $\mathcal{S}$.

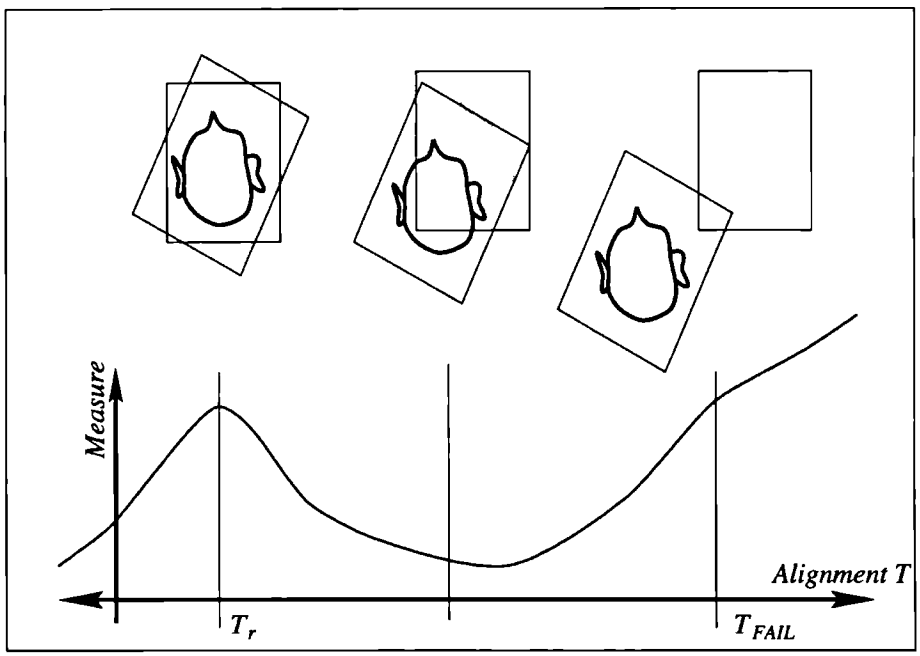


Figure 5.13: Image similarity is evaluated only from the region of overlap of two images and when derived from a small volume of overlap, any measure may give an invalid indication of alignment. Although there may be a distinct optimum at alignment (T_r) as the overlap falls there will always a point beyond which the measure provides an incorrect measure of alignment (T_{FAIL}).

The point at which an alignment measure provides an invalid indication of alignment will be determined by the extent of the images and also by their content. Because clinical

image volumes often have significantly different extents in each axis (in particular the z axis for transaxial slices) and can cut across the axis of the patient at very different angles, it is not possible to determine prior to registration the useful range of a similarity measure. This useful range is commonly referred to as the *capture range*. The practical definition of the capture range of a measure is also dependent on our method of optimisation. At best we can hope to recover registration from the subset of rigid transformations $\mathcal{T}_R$, for which the similarity measure is less optimal than its value at registration,

$$\mathcal{T}_C(T_r) = \{T | T \in \mathcal{T}_R, \mathcal{S}(m,n,T) < \mathcal{S}(m,n,T_r)\}. \tag{5.14}$$

This is an important factor in the choice of optimisation scheme. Stochastic schemes such as the Genetic Algorithm [42] have been proposed as an approach to finding the optimum of the 6D alignment parameter space [48, 51]. Initial experimentation with this approach highlighted the problem of choosing the extent of the starting population. In order to avoid choosing a registration estimate corresponding to a small volume of overlap, the spread of the GA population must be limited. This then reduces the ability of the Genetic Algorithm to avoid local optima. Such schemes are also expensive in terms of the number of evaluations required. In this application the computational cost of evaluating an estimate (the fitness of a single individual in the population) is expensive. As a result alternative optimisation approaches were investigated.

Starting Estimate

In general a valid starting estimate is provided by the imaging protocol itself, in that a specific region of interest, such as a tumour is imaged by both modalities. As such we can assume that the centre of the two imaged volumes are roughly aligned and there are no rotations. What we can do is to make the simple assumption that this single starting estimate (rather than a set of starting estimates) is within the capture range of the measure, and the final optimum lies up hill (or down hill, depending on the measure) from that estimate. The problem is then one of searching up hill from this starting estimate avoiding local optima on the way, as illustrated in figure 5.14.

5.5.1 Multi-Resolution Alignment

Multi-resolution techniques offer two major advantages during iterative registration:

- Increased Speed Through Computational Efficiency.
- Robustness to Local Optima.

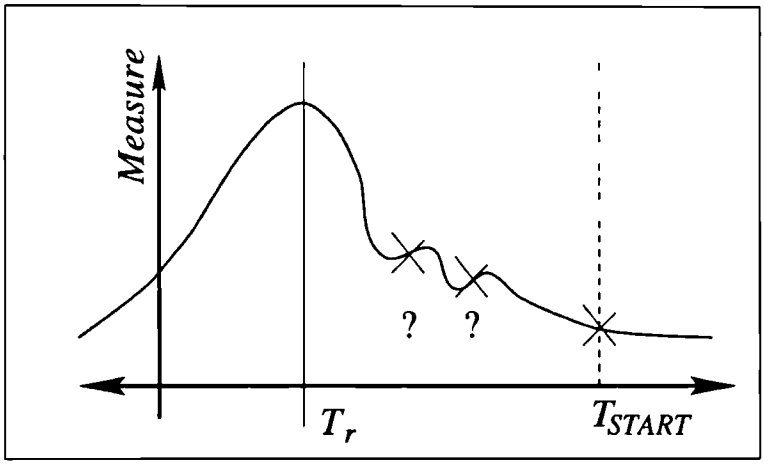


Figure 5.14: An illustration of a parameters space where the 'capture range' for even a simple uphill search is not easy to define because of local optima where the gradient may be reversed.

Multi-resolution techniques provide a direct method of speeding up the registration process [79, 105]. The advantage of using low resolution images is that they can be fully represented by a smaller number of measurement values. The alignment between images can therefore be evaluated from fewer corresponding measurements. This proves a significant factor when dealing with 3 dimensional images, where a reduction by a factor of 2 in the sample rate in each axis results in $\frac{1}{8}$ of the number of corresponding measurements.

As we lower the resolution of the images, smaller objects and fine structure are blurred from the images. The process of blurring of regions in the image corresponds to merging of smaller peaks with larger peaks in the joint histogram. As we continue to decrease the spatial resolution, the intensities in the images tend to the mean value in the image. In the joint probability distribution, larger and larger peaks, representing distinct regions merge, tending to a single sharp peak at the mean value of the two modalities. Averaging also decreases the uncertainty of noisy measurement values, which in turn reduces the spread around peaks in the joint probability distribution.

5.6 Multi-Resolution Optimisation

In order to improve the starting estimate a simple iterative optimisation scheme was used. Given a starting estimate T_0 described by six rigid transformation parameters, a similarity measure is evaluated for a set of 13 transformations $\mathcal{T}(T_0)$. These are the current starting estimate and the starting estimate with increments and decrements of each of the 3 translations ($\pm\delta t$) and 3 rotations ($\pm\delta\theta$).

5.6.1 Relating Changes in Translation and Rotation

Changes in rotations and translations can be related empirically by equating the displacement of a point 60mm from the origin (roughly at the skin surface for a typical brain image). So from Figure 5.15, given $x = 60mm$, for a given δt,

$$\frac{\delta t}{2} = 60 \sin\left(\frac{\delta\theta}{2}\right), \tag{5.15}$$

then,

$$\delta\theta = 2\sin^{-1}\left(\frac{\delta t}{120}\right) = 0.954° \tag{5.16}$$

This relationship between the rotations and translations was investigated experimentally and the value was not found to be critical to the final registration estimate.

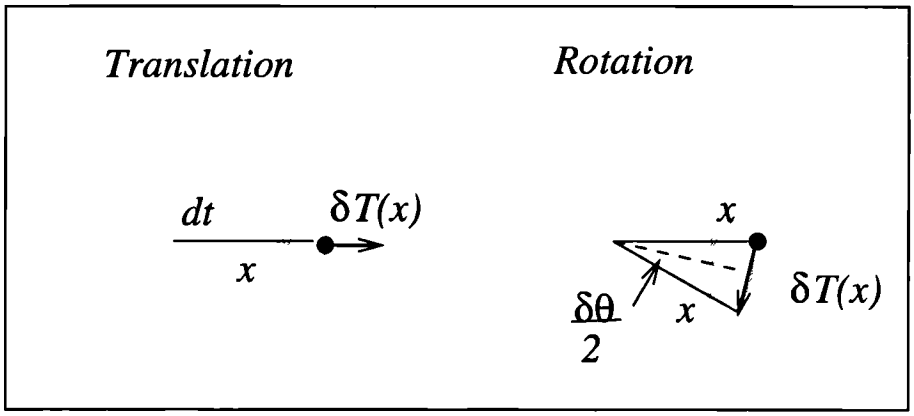


Figure 5.15: The affect of changes in translation and rotation on the displacement of a point at a given distance x from the origin.

We can then look for a better estimate of the registration transformation T_{n+1} with respect to measure $\mathcal{S}$ such that:

$$T_{n+1} = \max_{T \in \mathcal{T}(T_n)} \{\mathcal{S}(M(T), N(T))\}, \tag{5.17}$$

where $M(T)$ and $N(T)$ are the sets of image values occurring together for transformation estimate T. If $T_{n+1} \neq T_n$ then we can repeat the search with $\mathcal{T}(T_{n+1})$ until $T_{n+1} = T_n$, and no further improvement can be made. Following this the step size can be reduced and the search resumed.

The key factor in the approach is that the step sizes δt and $\delta\theta$ can be linked directly to the sampling resolution of the images from which the registration measure is being derived. By starting at low resolution and taking large step sizes we can efficiently find a rough estimate of the transformation. In addition, the low spatial resolution excludes small scale structures which may introduce local optima into the parameter

space. Because the accuracy of the estimate is limited by the spatial resolution of the images we must increase the evaluation resolution in order to further refine the estimate.

The problem with evaluating changes in the measure with respect to sub voxel displacements is that the response of the measure is dependent on the interpolation used in evaluating corresponding values in the two images. For computational efficiency tri-linear interpolation is used for this process. This introduces discontinuities in the estimate of the gradient of the interpolated values and therefore in the response of the joint probability distribution. As a result a gradient based search technique was avoided, in favour of this simple uphill search.

5.6.2 Initial and Final Step Size

Because of the nature of the parameter space between truncated image volumes, the choice of starting step size can be critical to the behaviour of the optimisation. Excessively large initial step-sizes mean that, even with a starting estimate lies inside the capture range, it is possible to step out of that capture range. Conversely too small an initial step size may result in the optimisation becoming trapped in a local optimum. For all results unless otherwise stated, optimisation was initiated at a data sampling resolution of $6mm$.

Experimentation with a small number of datasets indicated that the robustness of optimisation was increased in some cases by using intermediate levels in addition to those in the octree. This is effectively reducing the rate of reduction of the step size. For all the results presented here the sampling resolution was increased (and so step size decreased) by factors of $\sqrt{2}$ from $6mm$ to $3\sqrt{2}mm$, $3mm$, $1.5\sqrt{2}mm$ and finally $1.5mm$. At the highest resolution ($1.5mm$) the step size was reduced 8 times by a factor 2, resulting in a smallest translational step size of $\frac{1}{256}^{50}mm$ and rotational step size of $\frac{1.43}{256}^{\circ}$. This smallest step size was chosen experimentally as being appreciably smaller than the errors in the final registration estimate due to other factors. As will be seen in later experiments it is possible given these errors to select a larger final step size and therefore reduce the computational time required for registration.

5.7 Summary

This chapter was concerned with the practical evaluation and optimsation of image similarity measures. In evaluating image similarity it is important to consider the characteristics of volumetric image data, including its sampling, resolution, noise characteristics, extent and also padding regions.

To evaluate similarity between images we must compare discrete image measurements at an equivalent location (the interpolation problem) and over an equivalent region of material (spatial resolution). In interpolating and subsampling image values we must be careful to ensure that actual image measurements are not corrupted by padding regions within rectangular image data.

In evaluating the overlapping volume of particular regions of material with a discrete (binned) histogram, the effect of noise can be modeled as simply a reduction in our measure of this volume. This, given the unit of volume measured, then imposes a fundamental limit on the accuracy of our measure of alignment.

Chapter 6

An Experimental Comparison of Alignment Measures

6.1 Introduction

In chapter 3 we examined the ideas behind using voxel similarity to define a measure of alignment. Measures derived using information theory were identified as providing a generic approach to quantifying the alignment of different modalities. At the end of chapter 4 the differences in normalisation of these measures was highlighted as an important characteristic determining their behaviour as registration measures. In this chapter we shall examine these measures experimentally with the aim of identifying one measure which can be generically applied to MR-PET and MR-CT brain image alignment.

Essentially this chapter is concerned with one question: do the measures provide a meaningful indication of misalignment? We can answer this question by comparing their behaviour with respect to two important variables which govern the multi-modality alignment problem:

Image Content: How different or similar the images are.

Field of View: How much of the patient, and which region of the patient is imaged by both modalities.

Much work has been done in comparing registration measures by the evaluation and plotting of their values with respect to various axes of misregistration [21, 22, 48]. Because of the high dimensionality of the parameter space and the symmetries in the structure of the brain itself, the response of a measure to changes in one or two parameters alone can provide a poor indication of how the measures resolve the true 6 parameter rigid alignment solution. On the other hand we cannot realistically evaluate a significant portion of the full six dimensional space.

In practical terms, what we really want to know about the 6D parameter space is: given a starting estimate, can we find the optimum corresponding to registration? The approach which is taken in this chapter is to compare alignment measures when they are used within the multi-resolution optimisation scheme described in the previous chapter. All the measures may be evaluated from an estimate of the joint probability distribution. By using an identical approach to the evaluation of this distribution for a given transformation, together with the same approach to the optimisation of the measures, we can hope to compare their behaviour more precisely.

Firstly in this chapter, there is a brief review of the similarity measures that have been found suitable for MR-CT and MR-PET alignment. This includes examples of the structure of joint probability distributions for this type of image data. The remainder of the chapter then divides into two main experimental sections.

The first set of experiments apply the measures to align a database of MR-CT and MR-PET image pairs for which an accurate marker based estimate of the alignment is known. The database includes examples of images with and without large space occupying lesions and is thus representative of a range of MR-CT and MR-PET image content.

The second set of experiments look at the recovery of alignment from a range of starting estimates (determining the capture range) when we vary the field of view of the images. An image pair is selected and truncated by varying amounts axially and transaxially.

6.2 MR-CT and MR-PET Brain Image Alignment

6.2.1 Distributions of MR and CT Image Values

To illustrate the response to misalignment of the joint probability distribution between MR and CT values in the brain, distributions were evaluated between a pair of transaxially sliced 3D images at different levels of misregistration. The distributions were calculated from each overlapping voxel at registration and at a range of left-right translational misregistrations with respect to a fiducial marker based estimate. The images consisted of T1 ($TR = 650$, $TE = 15$), PD ($TR = 2550$, $TE = 20$) and T2 ($TR = 2550$, $TE = 90$) weighted MR sequences and a CT image of the brain of the same patient. The CT image has 28 $4mm$ slices of 512×512 $0.654mm^2$ pixels. The MR consists of 26 $4.128mm$ slices of 256×256 $1.265mm^2$ pixels.

Examples of these 2D distributions are shown in Figures 6.1,6.2 and 6.3 as surfaces whose height represents the percentage of the overlapping volume with a given pair of MR and CT values. The most common values (between 10 and 20 %) are those for air in

the two images, so to enable inspection of the smaller scale structures in the histogram, the height has been truncated. In addition a contour plot of the lower range of values has been added to highlight smaller changes in the distribution.

The main feature of all these distributions is the broad ridge of MR soft tissue values corresponding to a narrow range of CT soft tissue values. The shape of this peak in MR varies between the three types of MR sequence. One of the key features determining alignment is the location of the high CT values representing the skull with respect to the low MR values. As the images become misaligned, the high CT values over-lie higher MR values corresponding to soft tissue. This is reflected in the contour lines of all the distributions. An alternative view of this effect is given by a grey-scale image of these distributions as the two images are misaligned, which can be thresholded to highlight additional structure. An example is shown in Figure 6.4 which is evaluated for an MR-T1 and CT image pair where the images have been blurred to illustrate the affect of the different image boundaries on the distribution. These are seen as the 'hysteresis type' lines in the distribution. Here the true registration has been defined manually.

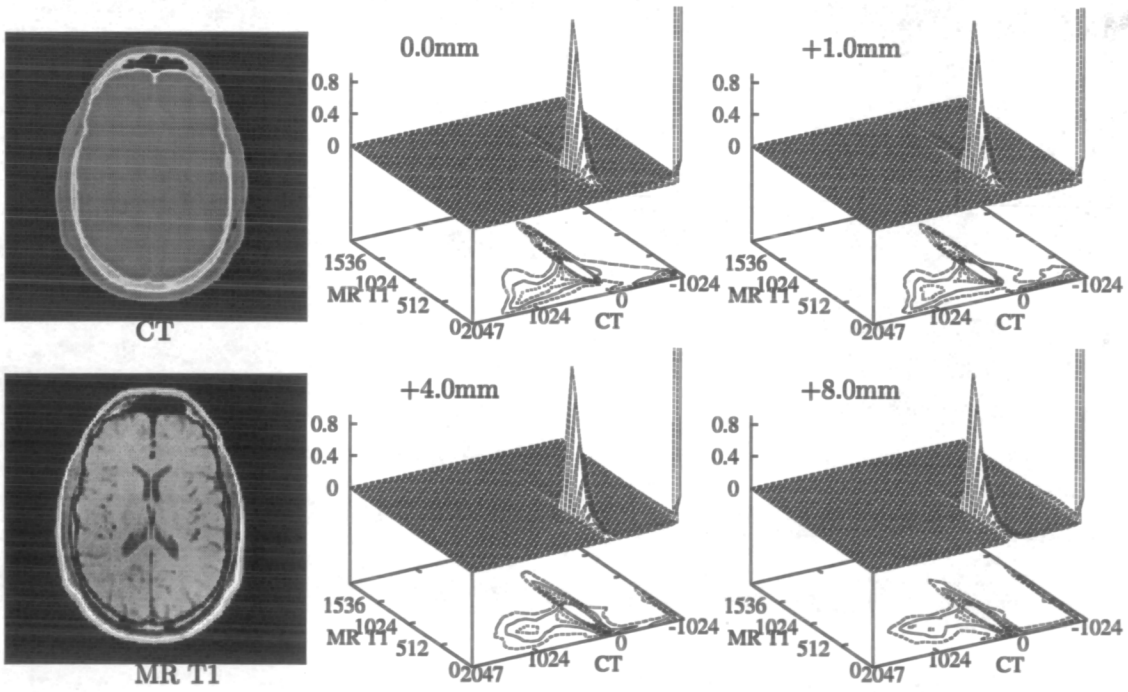


Figure 6.1: Examples of joint probability distributions formed between a T1 weighted MR (MR modulus values) and a CT (Hounsfield units) of the brain, evaluated at varying left-right translational misalignments.

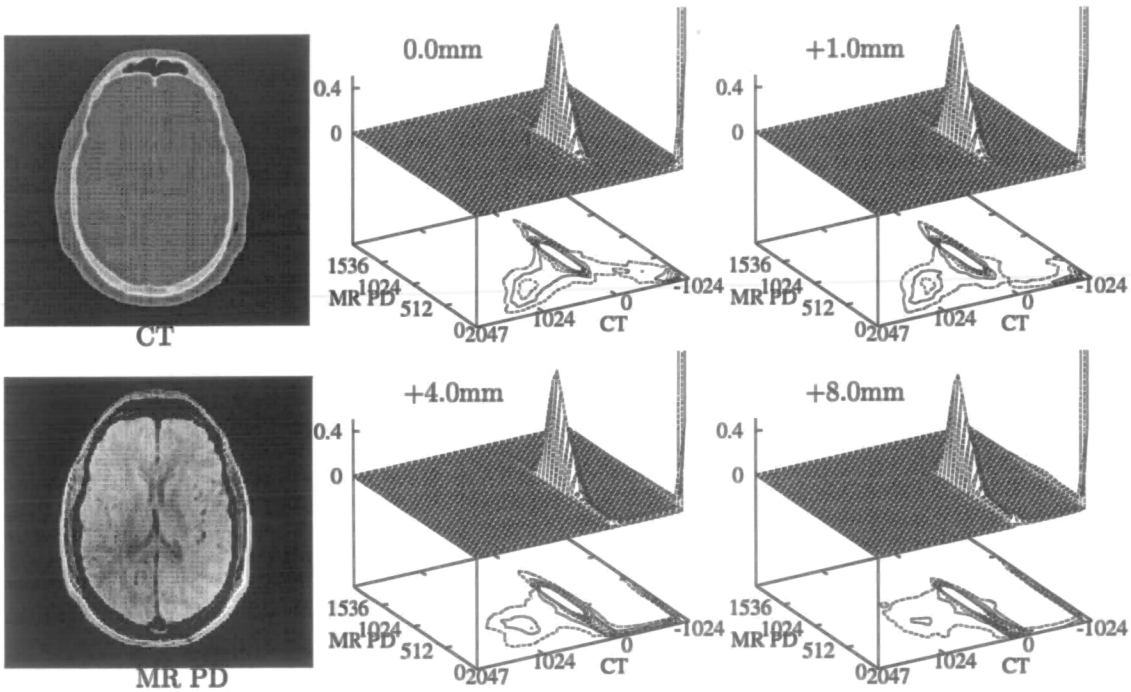


Figure 6.2: Examples of joint probability distributions formed between a proton density weighted MR (MR modulus values) and a CT (Hounsfield units) of the brain, evaluated at varying left-right translational misalignments.

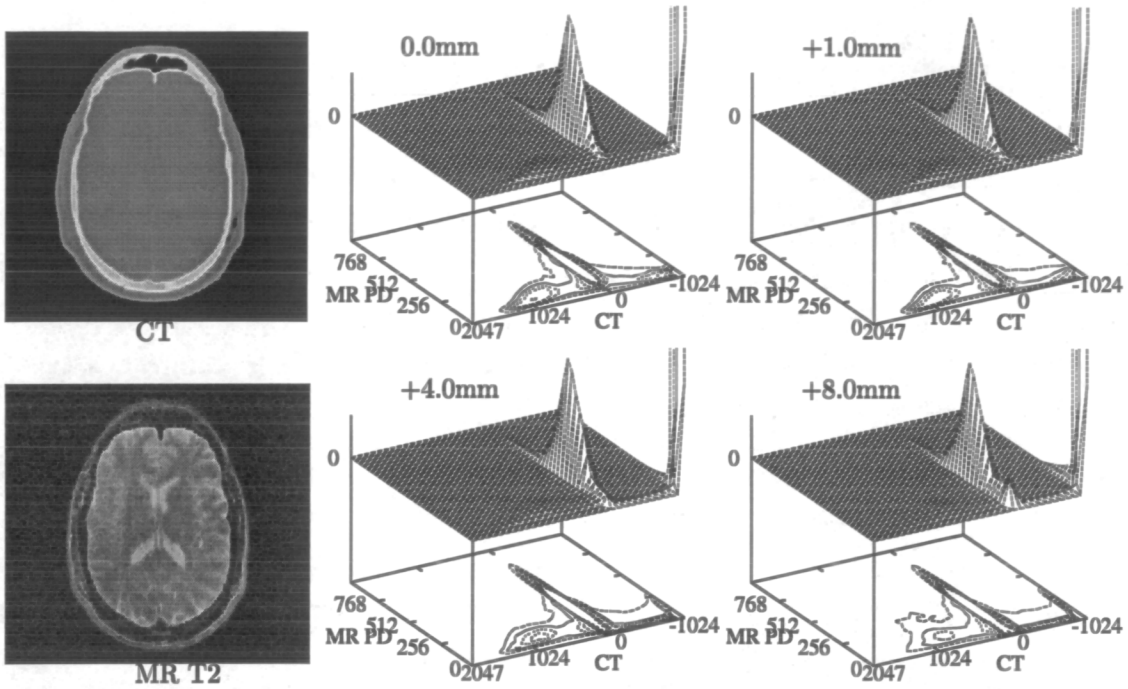


Figure 6.3: Examples of joint probability distributions formed between a T2 weighted MR (MR modulus values) and a CT (Hounsfield units) of the brain, evaluated at varying left-right translational misalignments.

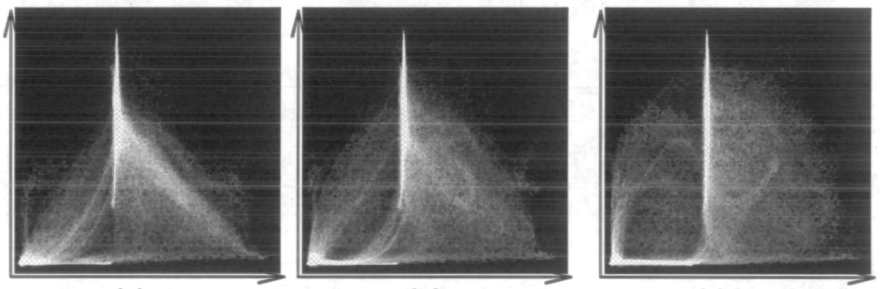


Figure 6.4: 2D histograms created from a CT (horizontal) and MR-T1 (vertical) image pair (a) manually registered, and (b) misregistered by translation along the cranio-caudal axis by 2 voxels, and (c) misregistered by 5 voxels. The intensity of each pixel $\{m, c\}$ in the histogram represents the number of voxels in the overlapping region of the two images that have a MR intensity m and a CT intensity c. The brighter the pixel, the more voxels in the images with that pair of values.

6.2.2 Distributions of MR and PET Image Values

We can repeat the same process of plotting distributions for MR and PET values as we used for MR-CT. Here we use the same MR images, but with a PET image of the same patient. Examples of the resulting 2D distributions are shown in Figures 6.5,6.6 and 6.7 as a surface whose height represents the percentage of the overlapping volume with a given pair of MR and PET values. As with MR and CT, the height has been truncated to show lower probability details.

The resulting joint histograms consist of a main central peak in the MR axis at around 700 in T1,T2 and PD and around 500 in the PET axis. This corresponds to the main grey and white matter brain tissue. Because of the noise and spatial resolution of these images there is poor distinction between grey and white matter peaks in the histogram. There is a peak or ridge in the MR (depending on the MR sequence) of values corresponding to approximately zero in PET. This corresponds to scalp, eyes and other non-brain tissues in the head in which there is little or no tracer uptake.

The relatively low signal to noise level of the PET image results in broad peaks in the PET axis. In addition, the significantly broad point spread function of the PET image (approximately $15mm$) results in partial volume values occurring between the main brain tissue peak and the air regions at registration. As for MR-CT, we can also form a grey-scale image of these distributions as the two images are misaligned, which can be thresholded to highlight additional structure. An example is shown in Figure 6.8 which is evaluated for an MR-T1 and PET image pair.

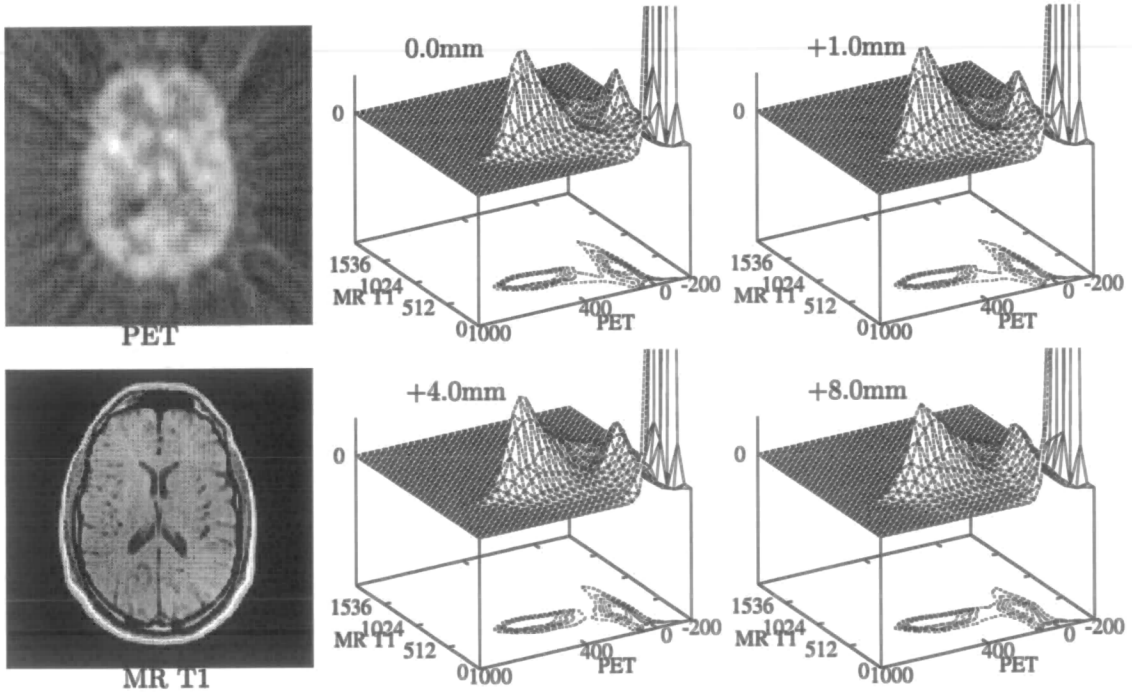


Figure 6.5: Examples of joint probability distributions formed between a T1 weighted MR and a PET image of a patient, evaluated at varying left-right translational misalignments

6.2.3 Measures

The experiments described in this chapter are intended to look at the problem of aligning MR with both CT and PET images of the brain. In these experiments we are primarily going to compare the three information theory derived measures,

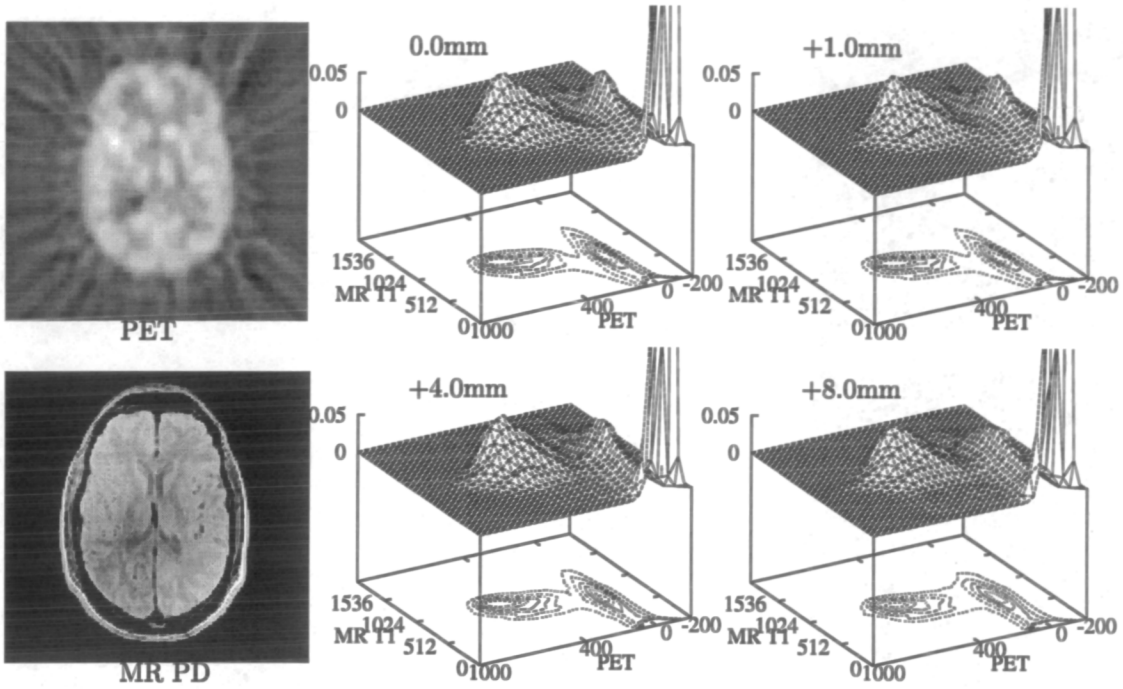


Figure 6.6: Examples of joint probability distributions formed between a proton density MR and a PET image of a patient, evaluated at varying left-right translational misalignments.

$H(M,C)$ **and** $H(M,N)$ Joint Entropy between MR-CT or MR-PET

$I(M;C)$ **and** $I(M;N)$ Mutual Information between MR-CT or MR-PET

$Y(M;C)$ **and** $Y(M;N)$ Normalised Mutual Information between MR-CT or MR-PET

when applied to both MR-CT and MR-PET alignment. These three represent the alternative approaches to relating the information content of the two images within their region of overlap. As a comparison, results using some of the statistical measures are included. These include the corresponding variance measures for both modality combinations:

$\sigma(M)$ Minimisation of MR variance for both MR-PET and MR-CT.

$\sigma(N)$ **and** $\sigma(C)$ Minimisation of CT or PET variance.

Finally we also look at correlation measures:

$\Gamma(M,C)$ Correlation Coefficient of MR with re-mapped CT soft tissue values

$\Gamma(M,N)$ Direct Correlation Coefficient of MR with PET FDG values

For MR-PET we can evaluate correlation directly, while for MR-CT we can apply the re-mapping approach used by Van Den Elsen and described in section 3.3.2. CT soft

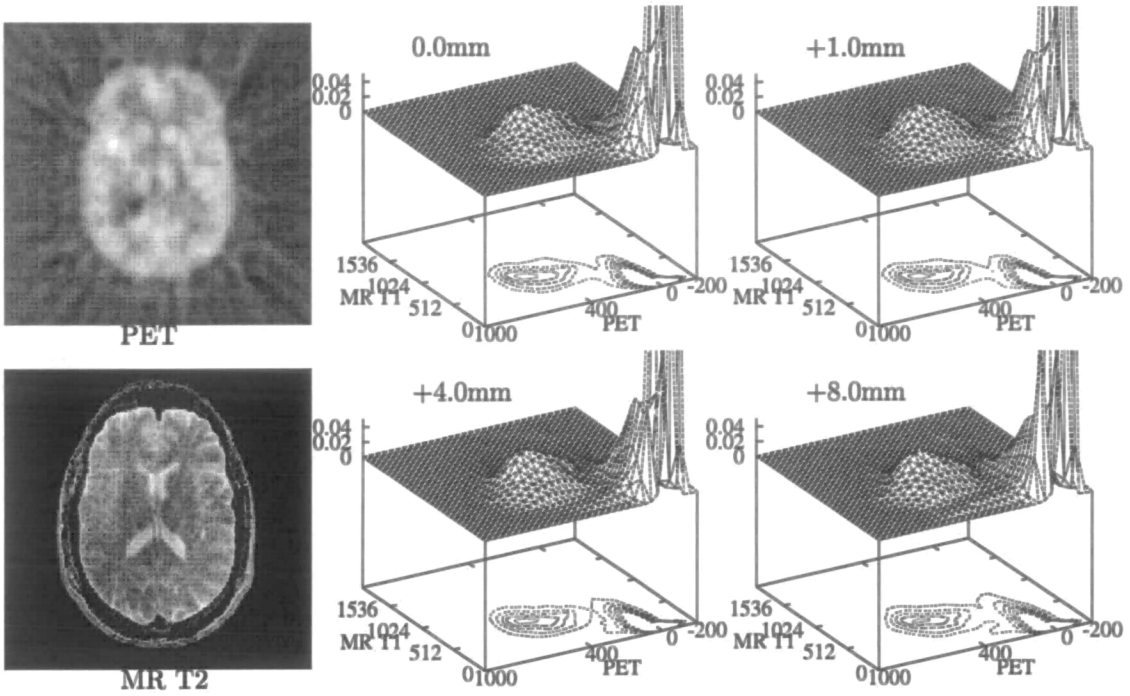


Figure 6.7: Examples of joint probability distributions formed between a T2 weighted MR and a PET image of a patient, evaluated at varying left-right translational misalignments.

tissue re-mapping was found to provide significantly better behaviour than the bone re-mapping approach for this data, and so it is this measure which is included in this comparison.

Both the correlation and variance measures can be derived directly from the same discrete estimate of the joint histogram as the entropy measures. All measures can then also be optimised using the same multi-resolution scheme described in the previous chapter.

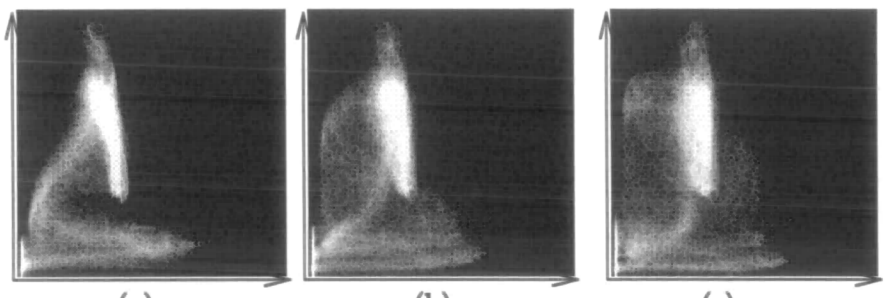


Figure 6.8: 2D histograms created from a PET (horizontal) and MR-T1 (vertical) image pair (a) Manually registered, (b) misregistered by translation along the cranio-caudal axis by 4 MR voxels, and (c) by 10 MR voxels. The intensity of each pixel (m, n) in the histogram represents the number of voxels in the overlapping region of the two images that have a MR intensity m and a PET intensity n. The brighter the pixel, the more voxels in the images with that pair of values.

6.3 The Vanderbilt Image Registration Database

6.3.1 An Independent Registration Estimate

Given that we have different measures of alignment, we would like to be able to compare estimates provided by optimisation of these measures to a 'ground truth'. One plausible approach is to create synthetic image data containing simulated structures delineated in different modalities [61]. The validity of such an approach is limited by the accuracy of the simulation of real clinical image structure. The small scale similarities and differences in anatomical boundaries, and the measurements exhibited by tissues in a patient will determine the behaviour of the system. Creating an accurate simulation of this structure that contains all the anatomical features which can influence the similarity measures is beyond the capabilities of current anatomical models. Using synthetic image data as a gold standard can therefore only give an indication of the behaviour of different techniques. A valid test of accuracy and robustness must employ real image data.

The problem with using real image data is then one of defining an independent estimate of rigid alignment between two images of a patient. This turns out to be a very difficult problem. Previous approaches to defining an estimate for images of the head have employed markers visible in two modalities attached to the skin [26], a stereotactic frame, a frame held in the teeth [90] or a multi-modality compatible head holder [30]. These, in general, have provided estimates with limited accuracy due to the movement of the markers with respect to the rigid internal structures of the skull.

One of the most accurate approaches so far has been that developed at Vanderbilt University [103], and was used as part of their Phase I Retrospective Registration Evaluation Project (RREP). It is part of this database of images which will be used for the experiments in this chapter. The method employs 4 fiducial markers rigidly fixed to the skull in conjunction with a number of additional techniques to ensure the accuracy of the global rigid alignment estimate. These included

- The use of a stereotactic frame to prevent patient motion during an acquisition.
- Independent estimates of MR scaling parameters for the images, derived from points on the stereotactic frame.
- MR static field inhomogeneity correction by the use of dual acquisitions with opposing readout directions [16].

Because of the nature of the protocol, which makes use of a stereotactic frame, all the acquisitions have a large field of view covering a significant portion of the brain, and have similarly orientated slice planes. The images also have a relatively large slice

thickness (typically 4.0mm), compared to the majority of brain images acquired at our site.

6.3.2 Image Data

The complete database of images in the RREP study consisted of both rectified and non-rectified images. The experiments in this section were only carried out using those image pairs for which geometric rectification had been applied. This ensures the most accurate gold standard for the rigid registration. This image data consists of CT images acquired for each patient along with T1-weighted, T2-weighted and Proton density (PD) weighted MR images. All the data was nominally axial and acquired with a stereotactic frame. This was then removed from the images as part of the blinded study, and replaced by noise signal simulating air. For the experiments in this thesis, only patient data sets for which all three MR sequences were available in geometrically rectified form were included. This means that patient sets 1,2,5 and 7 were used for the MR-PET experiments and patient sets 1,2,3,4,5 and 7 were used for the MR-CT experiments.

The CT images had a sampling resolution of $0.65 \times 0.65 \times 4.0mm$ with 512×512 pixels in plane, and between 28 and 34 slices. The MR data had a nominal sampling resolution of $1.25mm \times 1.25mm \times 4.0mm$ with 256×256 pixels in plane and between 20 and 26 slices. These images were corrected to account for MR scaling errors, resulting in slightly modified voxel dimensions. A number of the images contained significant volumes of space occupying lesions, which were visible on one or more of the acquisitions.

6.4 Assessing Registration Quality

This section describes the two main approaches that will be used in this thesis to evaluate and express the accuracy of a registration estimate. First there is a discussion of the approach used in expressing the difference between an estimate and an independent gold standard. Secondly there is a brief description of the basic tools and protocol used for visually assessing alignment accuracy where there is no independent estimate of alignment.

6.4.1 Expressing Accuracy Relative to a Gold Standard

MR-CT Alignment Accuracy: MIPVD

There are a number of alternative approaches to comparing the quality of transformation estimates. Firstly, we can directly compare the 6 rigid parameters of the estimates which gives an indication of the direction of any misalignment. A more clinically relevant measure of the accuracy of an estimate is the displacement of anatomical locations

within the patient from their 'correct' location. Evaluating registration accuracy over the full imaged volume can provide a misleading estimate of registration accuracy for the actual patient tissues. This is particularly important for image data which has a large field of view, where only a small portion in the centre of the image is occupied by the patient. The approach used by some workers is to select a set of relevant anatomical landmarks at which to measure the point displacement, as used in the Vanderbilt study [103]. This approach is not easily applicable to the whole of our database of clinical image data (which will be described in the following chapters), because of the great range in anatomical coverage and quality of the images. We would ideally like a simpler definition of anatomical regions which can be applied to this data.

The approach used in this work is to extract a set of imaged points V_C falling within the patient $V_P \subseteq V_C$ by interactively defining a simple thresholding of CT image values to exclude air voxels. The voxel displacements were then evaluated only for the subset of patient voxels which also overlapped with the imaged MR volume V_M,

$$\mathcal{V}_P = V_P \cap T_r(V_M), \tag{6.1}$$

for a reference registration estimate T_r. This avoids problems (particularly with later data) where the CT and MR cover substantially different volumes.

For this data, a CT minimum threshold of -200 (Hounsfield units) was used, which excluded regions of air. This region is illustrated for the CT for patient 1 in Figure 6.9.

For each of the discrete voxel locations $\mathbf{x} \in \mathcal{V}_P$, as illustrated in Figure 6.10, the point displacements,

$$\mathbf{D}_P = T_s(\mathbf{x}) - T_r(\mathbf{x}). \tag{6.2}$$

between an estimate T_s and the reference estimate T_r were evaluated. The mean $|\bar{\mathbf{D}}_P|$ voxel displacement gives an overall indication of the quality of an estimate for patient volume which is imaged by both modalities. This is termed the mean imaged patient voxel displacement (MIPVD).

MR-PET Alignment Accuracy

Essentially an identical approach to that used for CT transformation accuracy can be applied to PET images. In order to define a set of patient voxels V_p, again a simple threshold of, in this case, the PET image was used. Because of the different properties of each PET image, individual thresholds defining patient voxels were set interactively for each image. These thresholds essentially excluded all areas of non-brain voxels from the assessment of patient displacement, but because of the increased noise of the PET images, some artefactual regions are sometimes included, as illustrated in Figure 6.11.

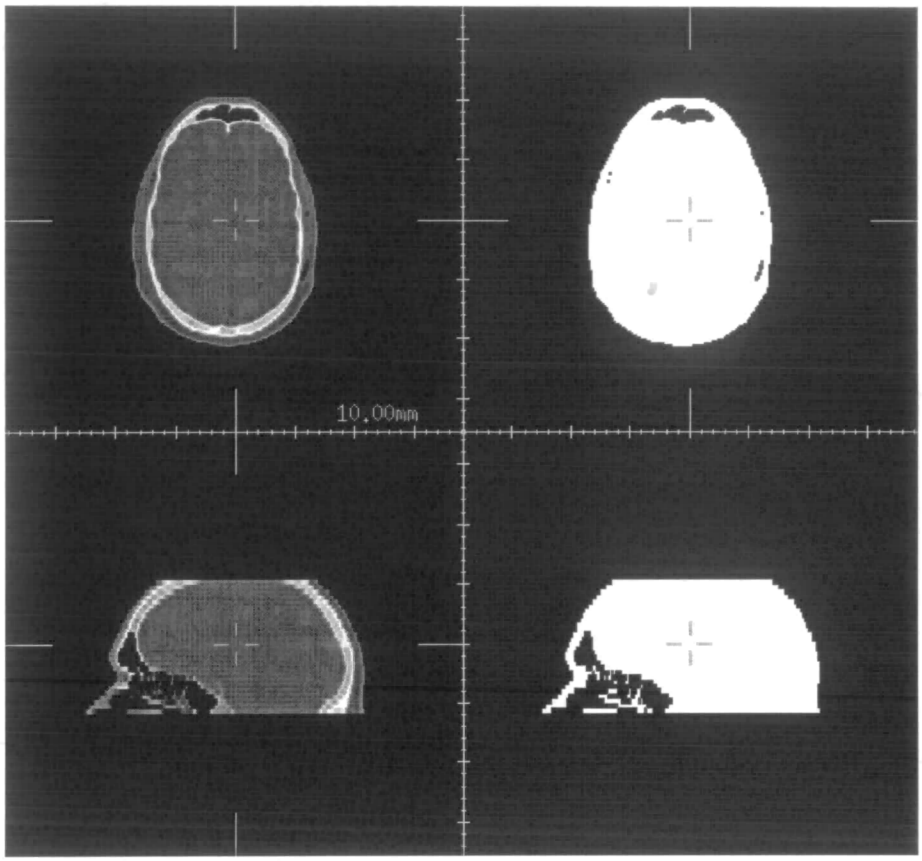


Figure 6.9: Example axial (top) and sagittal (bottom) slices through the CT volume for patient 1 (left) with thresholded patient voxels V_P used for accuracy validation shown (right).

6.4.2 Visual Evaluation of Image Alignment

In many cases when using clinically acquired data in experiments it is not possible to have an independent estimate of registration. In such cases it is still possible to assess the quality of an automated estimate visually using appropriate visualisation tools. An important factor in enabling much of the research carried out in this thesis has been the continued development of an interactive registration visualisation tool called *rview*. This was developed as part of this thesis work in C++/MOTIF under UNIX to provide a way of interactively examining the alignment of two 3D images.

3D volumes are displayed as orthogonal cuts at a user defined location in space. Accurate evaluation of alignment is made possible by the overlay of an iso-intensity contour from one modality onto a second in each of these planes. This mode of display is particularly suited to the visualisation of MR-PET and MR-CT alignment. Interactive selection of threshold level to define a contour in either CT or PET provides a simple way for the user to select and display corresponding features overlayed onto the MR.

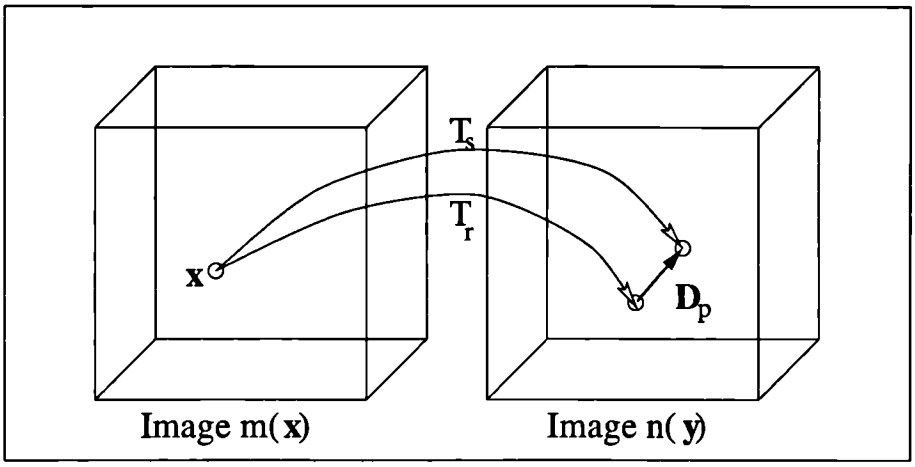


Figure 6.10: Evaluation of point displacements between a reference transformation T_r and an estimate T_e.

For MR-PET alignment a contour level in PET is selected interactively to correspond to brain surface. By using the views in the three planes simultaneously it is possible to define a level which delineates a surface corresponding to the brain delineated in MR as illustrated in figure 6.12. When overlayed on the MR for a given transformation estimate this provides a sensitive indication of misalignment.

A similar approach is taken to assessing the alignment of MR and CT. Here a contour level is defined in the CT to delineate bone surface as illustrated in figure 6.13. When this contour is overlayed onto MR this again provides a good indication of the quality of image alignment.

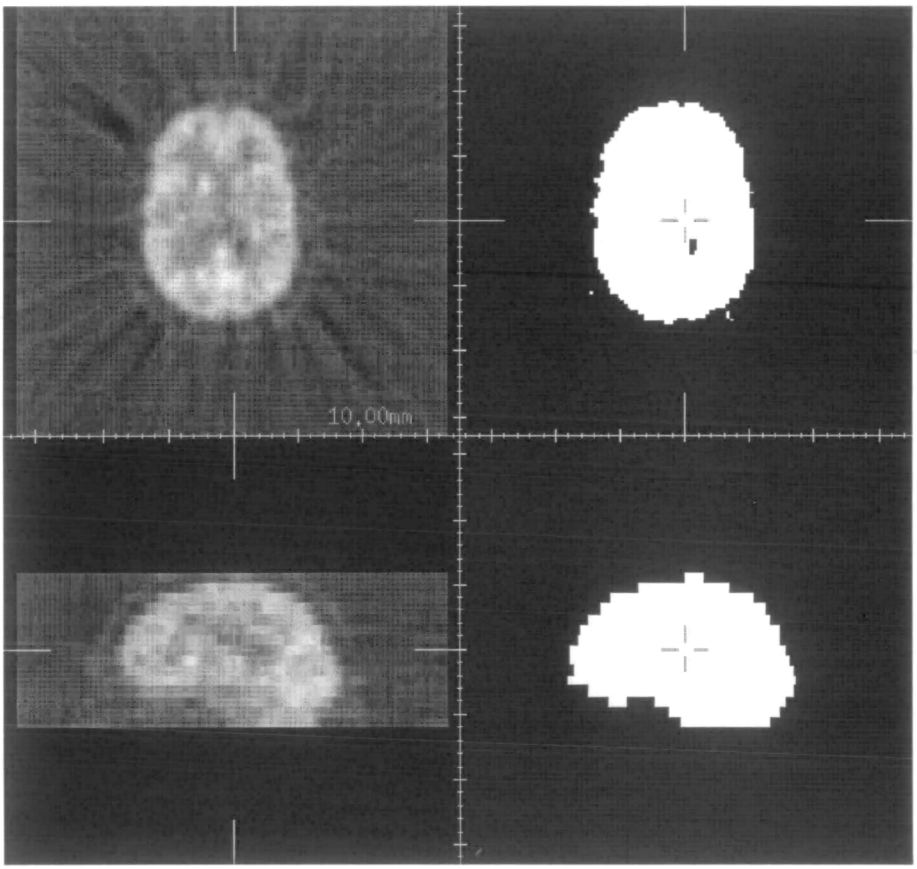


Figure 6.11: Example slices through image set for patient 1 showing thresholded region selected for accuracy measurements.

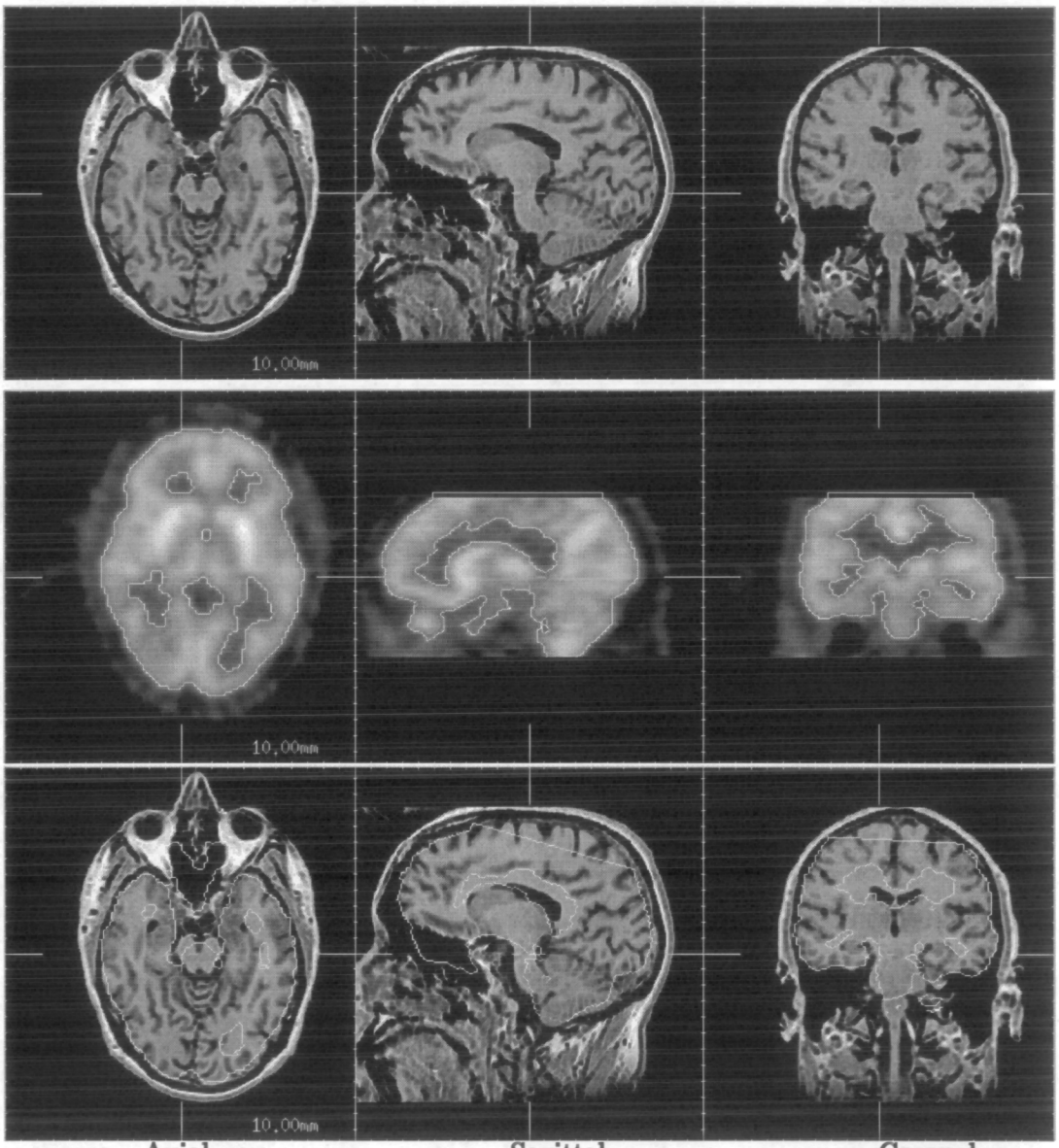


Figure 6.12: Inspection of an acceptable manual point based MR-PET registration estimate by interactive selection of a PET contour level (middle) to delineate brain surface. This is then overlayed using the transformation estimate onto the MR (bottom.)

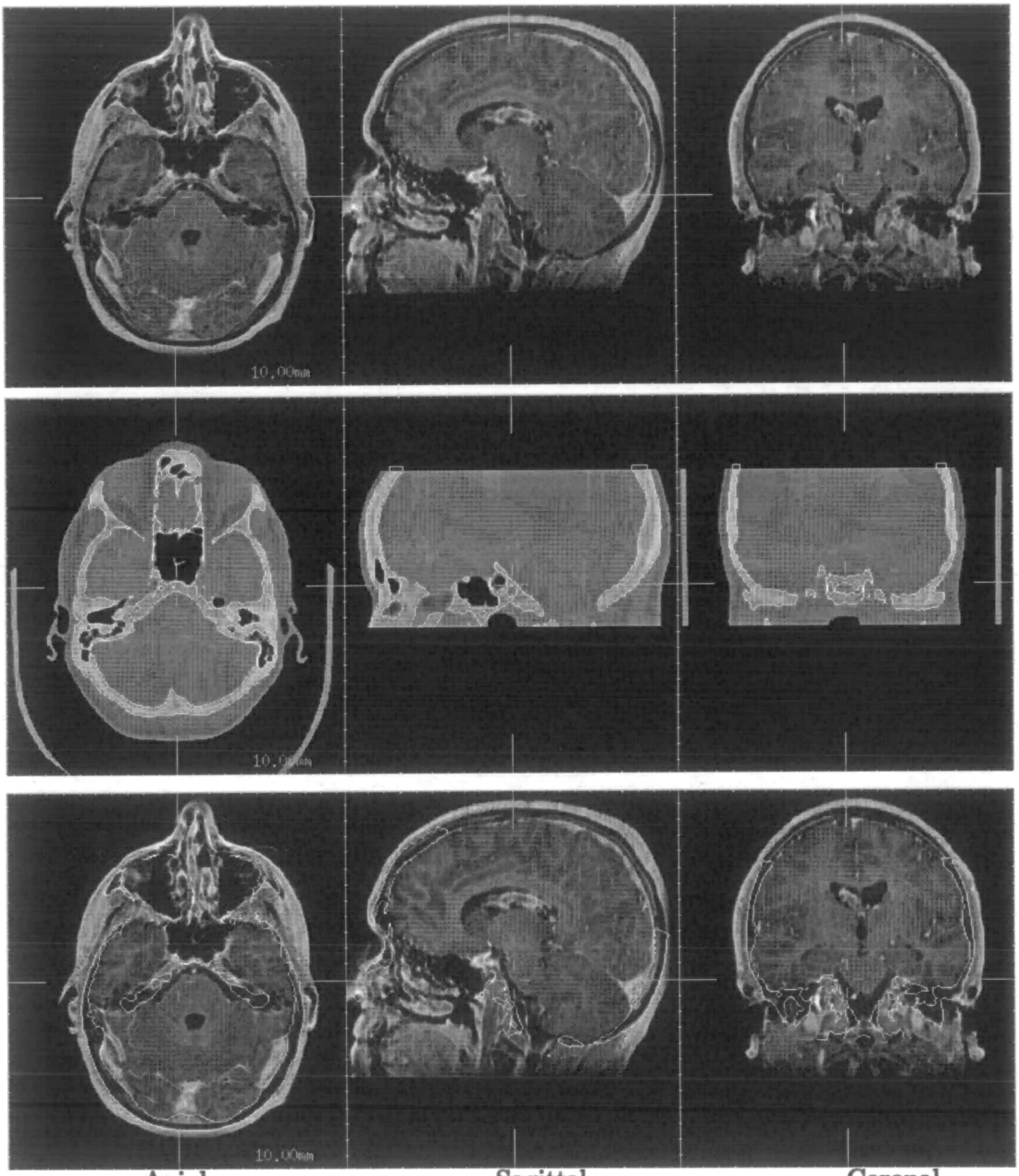


Figure 6.13: Inspection of a manual point based MR-CT registration estimate by interactive selection of a CT contour level (middle) to delineate bone surface. This is then overlayed using the transformation estimate onto the MR (bottom.)

6.5 Robustness and Accuracy of Registration with Respect to Image Content

In this first set of experiments we are going to determine whether the different similarity measures can provide a useful measure of alignment between a range of MR-CT and MR-PET images. To do this we are going to use them to try and recover the alignment of images containing typical clinical content. We need to know if the parameter space is such that we can find the optimum corresponding to alignment (we can call this the capture range), and also if this optimum does actually correspond to correct alignment (accuracy).

6.5.1 Experimental Design

The starting point for the investigation of the behaviour of the measures is to attempt to recover the alignment of each of the image pairs directly using the different measures. The same default starting estimate as described in section 5.5 was used for each image pair and each measure. This provides an indication of the capability to automatically recover alignment without any prior knowledge of the nature of misalignment. In addition the same optimisation parameters as described in section 5.6 were used to carry out the registration.

6.5.2 Results: MR-CT

The bar-charts in Figure 6.14 and 6.16 summarise the quality of the estimates for each of the image pairs with respect to the Marker based estimate. It is obvious from these that there are a number of significant failures for some of the measures

Minimisation of corresponding MR variance, $\sigma(M)$, (Figure 6.14) fails for all the images, providing estimates far from the marker estimate. An example failure is shown in Figure 6.15, illustrated, as described in section 6.4.2, by the overlay of a bone contour from the CT onto the MR. This confirms the expectation that the mapping from CT to MR is too far from being a simple one-to-one mapping (see section 3.4.1) for this statistic to provide a robust measure of alignment. Many of the misalignments contain large translational components, presumably reducing the overlap of the patient tissues in the two images. In contrast, minimisation of CT variance provides estimates close to the marker estimate for all the image pairs.

CT soft tissue correlation provides estimates close to the marker estimate, except for two significant failures (The MR-T1 image pairs 4 and 7). Examination of the results show that optimisation has found a visually plausible local optima. Figure 6.17 shows that the poor delineation of the skull boundary in much of the MR image (see sagittal

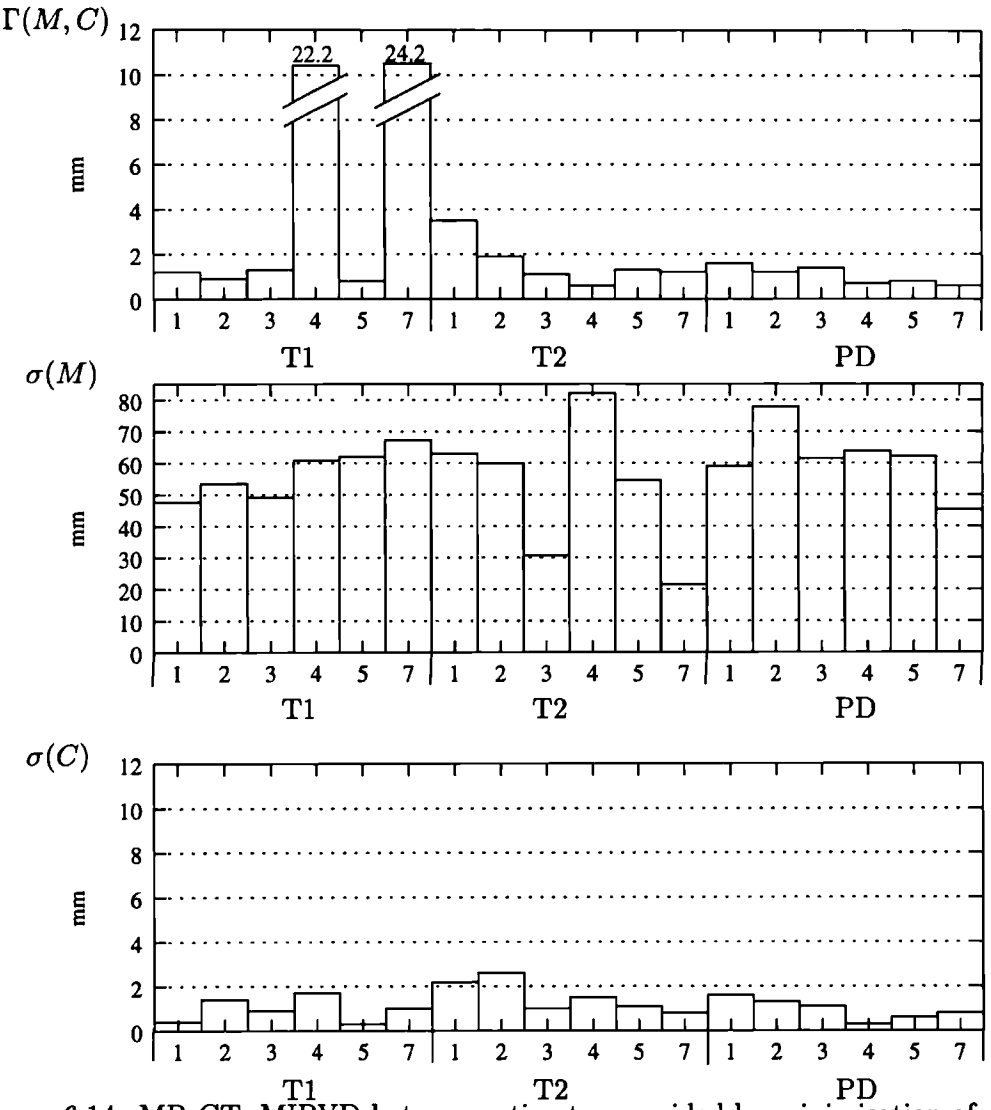


Figure 6.14: MR-CT: MIPVD between estimates provided by minimisation of variance and correlation of soft tissue measures, and the marker based estimate. (Note different vertical scale for the measure $\sigma(M)$.)

plane) may be a cause of this incorrect optima. In the re-mapped CT 'soft tissue' image, the bone of the skull will be clearly distinguished. This may then degrade or narrow the optima at or near true registration.

Moving on to the entropy measures, the estimates provided by minimising joint entropy show consistently good alignment, except for one of the patient image pairs (3 T1). Visual inspection of this result shows that the optimisation appears to have found a poor local optima of the measure (Figure 6.18), where the skull in CT is partially aligned as indicated by the white arrow.

Optimisation of mutual information from this starting estimate provides poor results with complete failures for about half of the image pairs. Inspection of the estimates (fig

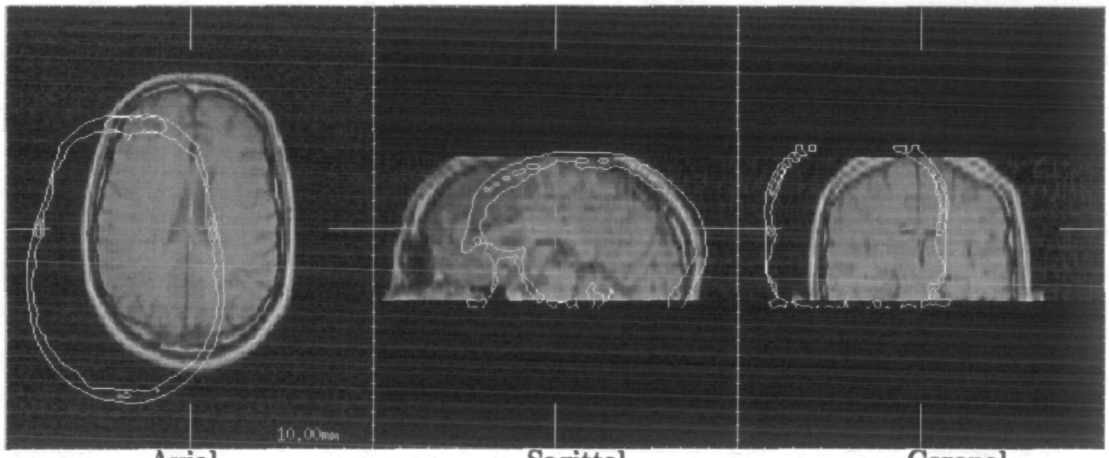


Figure 6.15: Slices through MR image volume for patient 2 illustrating registration failure resulting from the optimisation of $\sigma(M)$.

ure 6.19) indicates that the optimisation has found a visually plausible optima far from true registration, with many bone features roughly matching in the region of overlap. The rotational misalignment in the coronal plane confirms that the solution found has tended to reduce the quantity of 'background' voxels in the volume of overlap. This confirms the the findings of the simulation experiments in section 4.4.

Normalised mutual information provides the best overall results on this database of image pairs. A visually plausible estimate close to the marker estimate is provided for all image pairs.

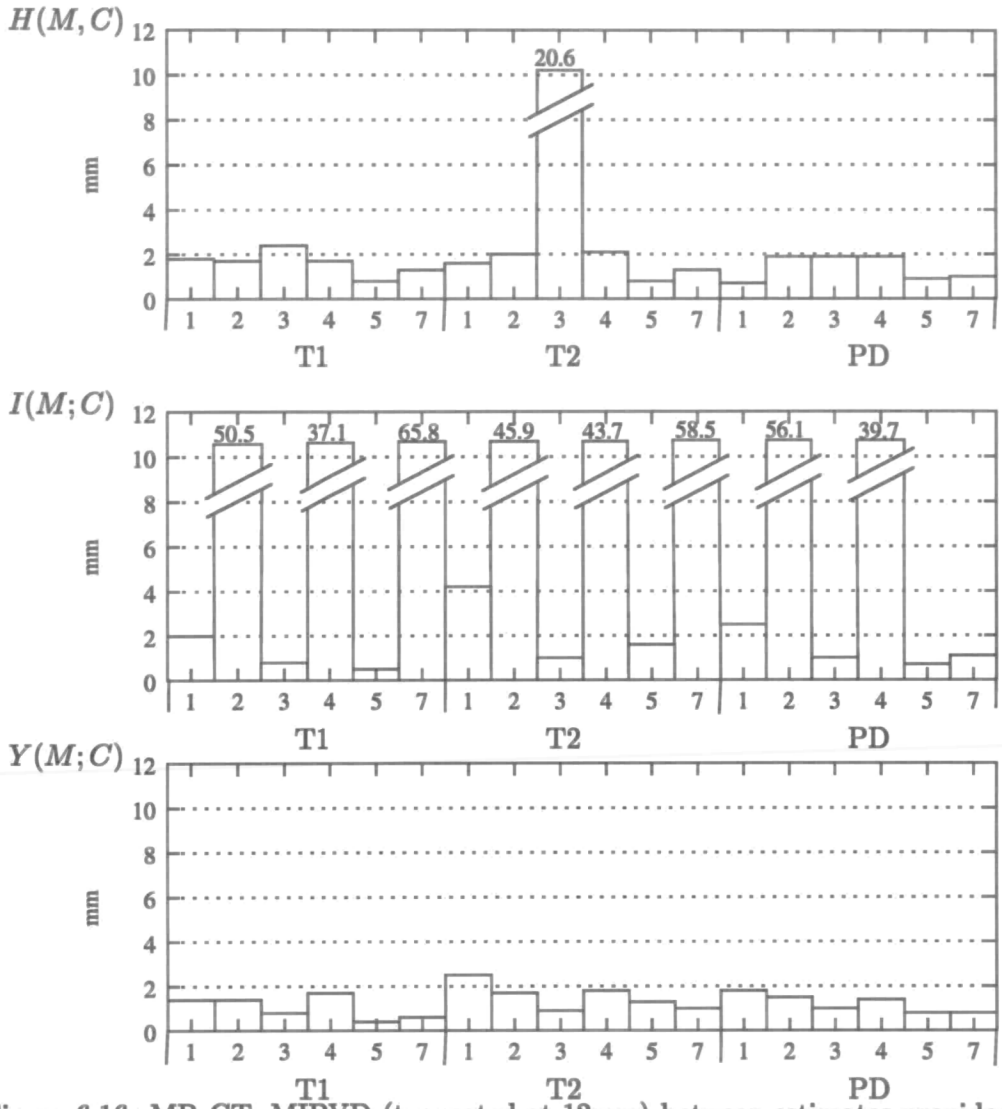


Figure 6.16: MR-CT: MIPVD (truncated at 12mm) between estimates provided by the three entropy measures and the marker based estimate.

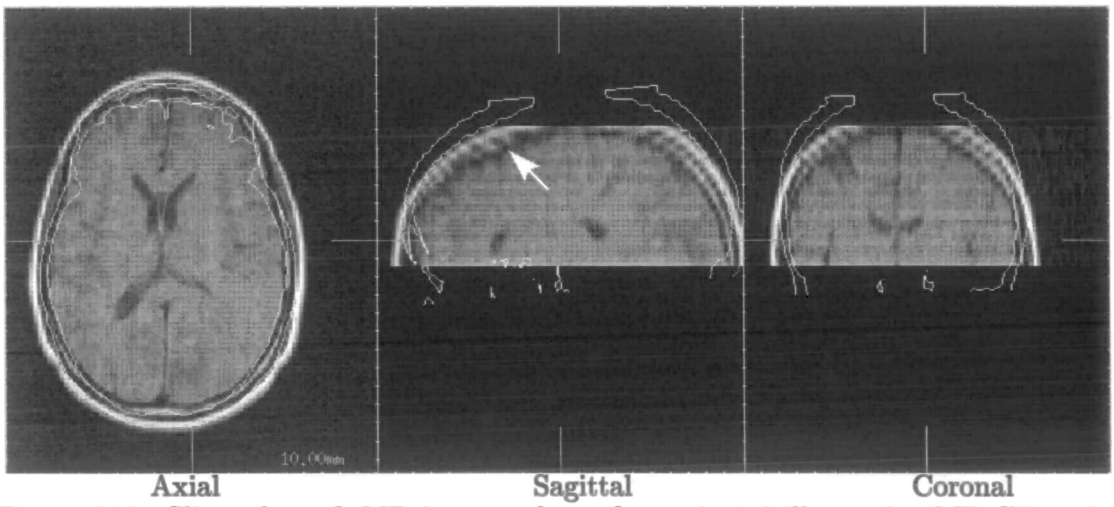


Figure 6.17: Slices through MR image volume for patient 4 illustrating MR-CT registration failure resulting from the optimisation of $\Gamma(M, C)$.

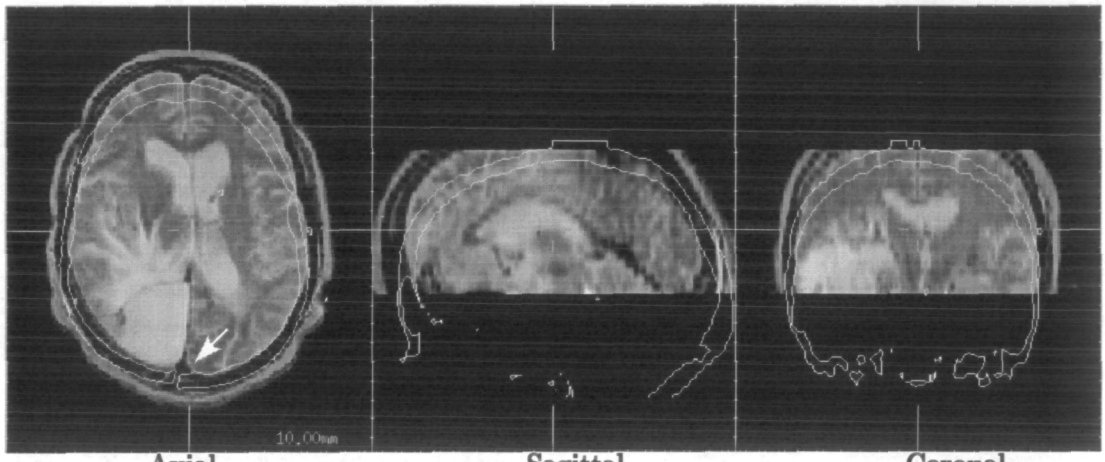


Figure 6.18: Slices through MR-T2 image volume for patient 3 illustrating registration failure resulting from the optimisation of $H(M, C)$.

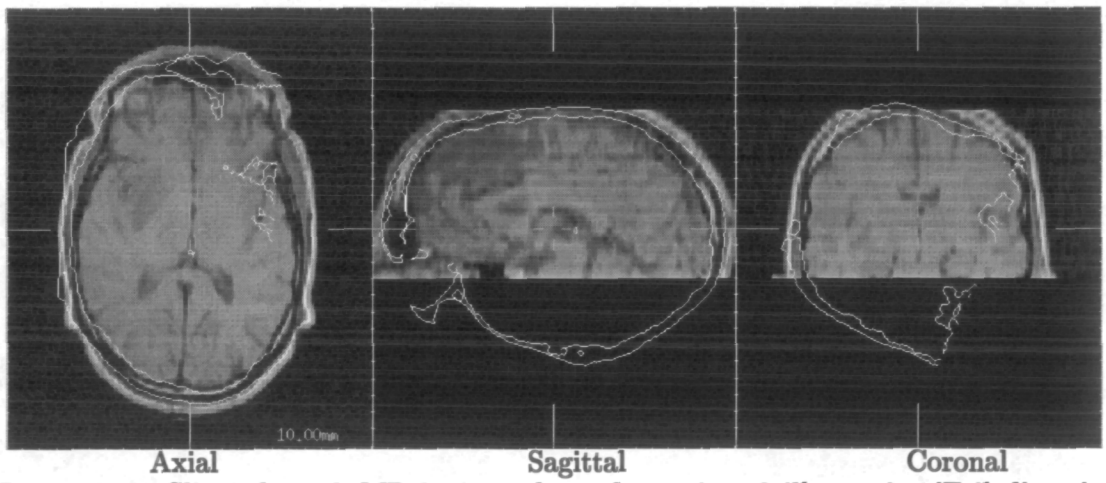


Figure 6.19: Slices through MR image volume for patient 2 illustrating 'Failed' registration estimate resulting from the optimisation of Mutual Information $I(M; C)$ (Note: in-plane transaxial field of view truncated in this display.)

6.5.3 Results: MR-PET

As with CT, the results are summarised by the bar-graphs shown in Figures 6.20 and 6.21, showing the mean imaged patient voxel displacement for each image pair from the marker estimate. Overall, all estimates are further from the marker estimate than was the case for MR-CT alignment. This reduction in accuracy is expected given the considerably lower spatial resolution of the PET imagery.

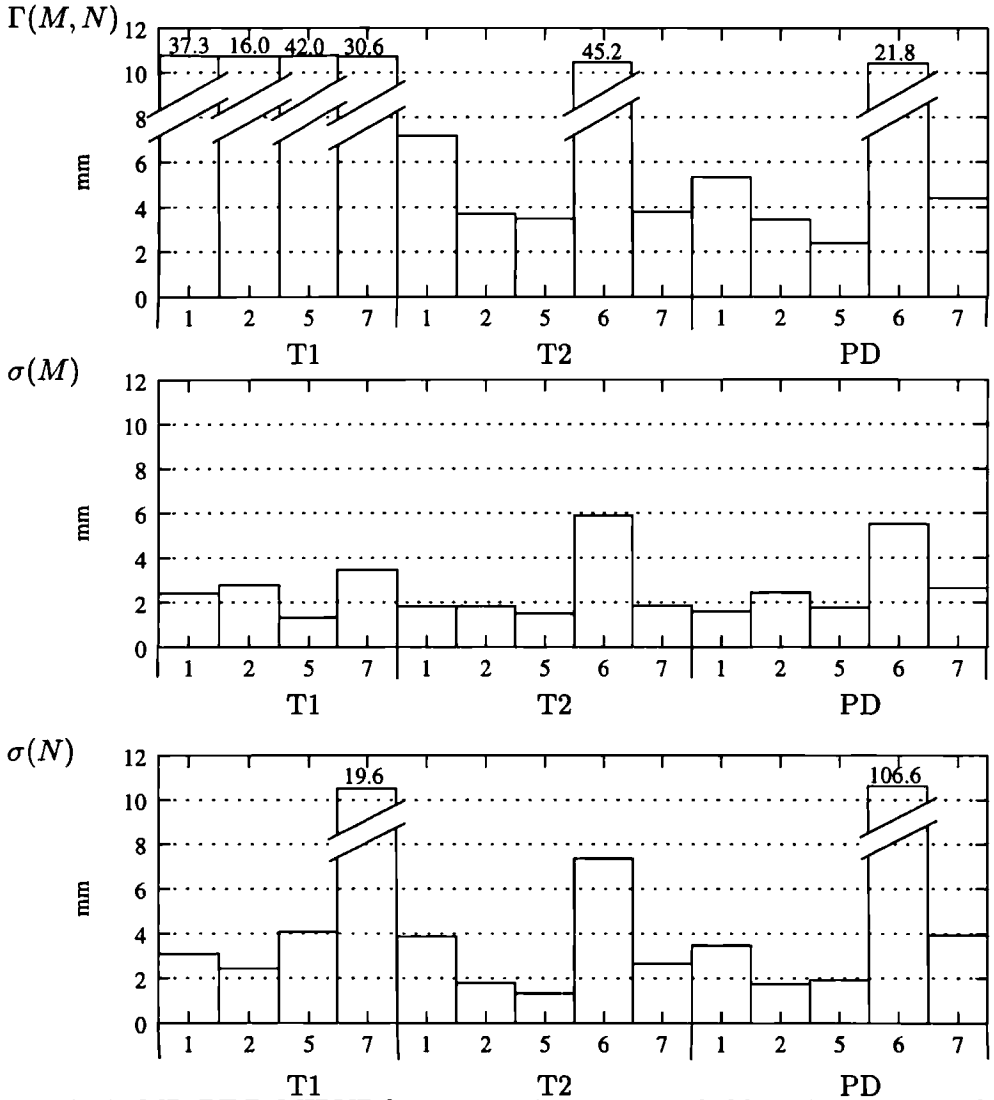


Figure 6.20: MR-PET: MIPVD between estimates provided by minimisation of variance and correlation coefficient measures, and the marker based estimate (note difference in vertical scale.)

Correlation, in particular proves a poor measure of registration for many of the image pairs in this database. In other work [86] it was found that correlation can provide a relatively robust measure on MR-T1 data acquired for functional brain studies. Here though the images contain significant space occupying lesions and include a range of

different MR sequences which reduce the linearity of the joint probability distribution at registration.

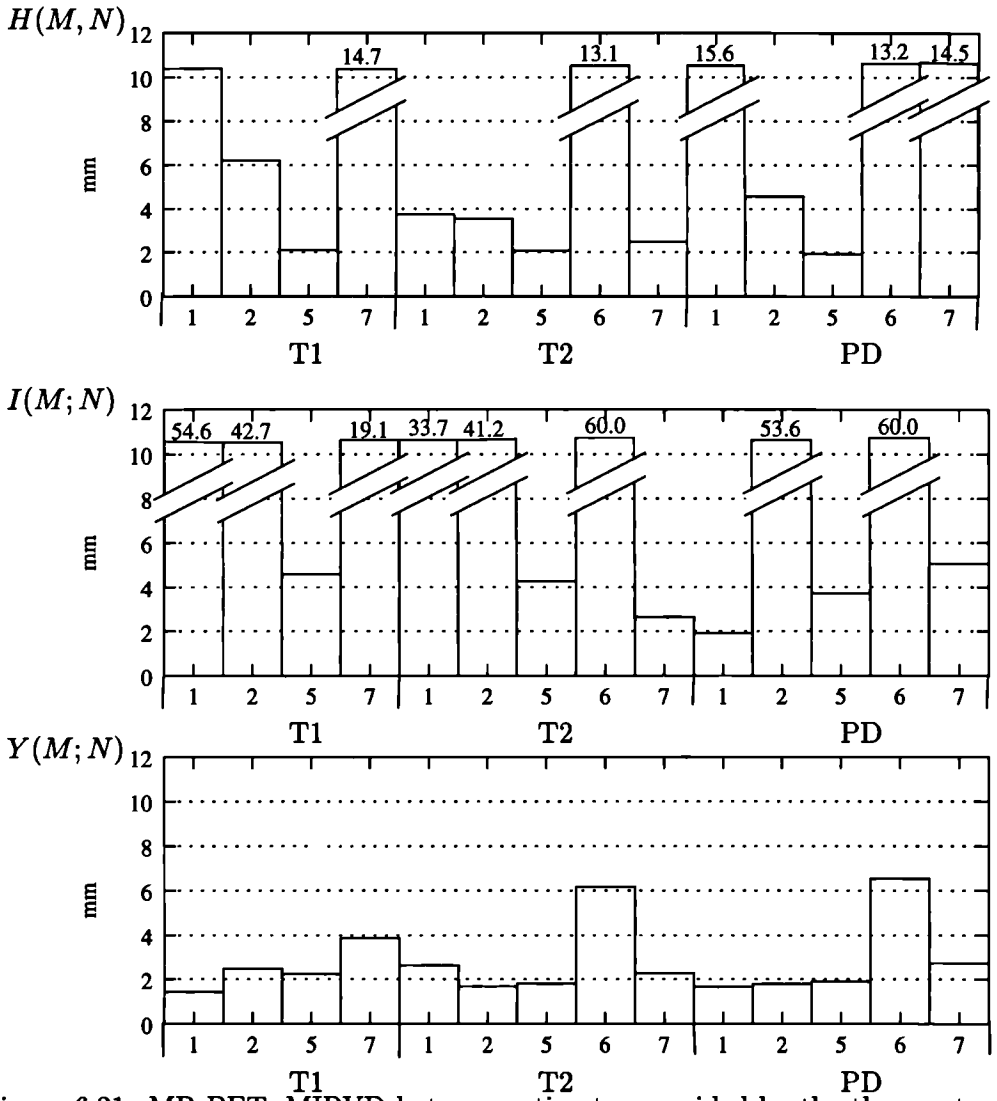


Figure 6.21: MR-PET: MIPVD between estimates provided by the three entropy measures and the marker based estimate.

Unlike MR-CT alignment, the two variance measures both provide estimates close to the marker estimate for the majority of image pairs. This may be because both MR and PET delineate grey and white matter in the brain. The minimisation of PET variance ($\sigma(N)$) provides poor estimates for three image pairs. The same image pairs also show appreciable misalignment errors with the MR variance ($\sigma(M)$) measure.

The estimates minimising joint entropy between MR and PET images (Figure 6.21) show some failures and some successes, while again those using mutual information are significantly worse, with similar types of mis-alignment found to those of the MR-CT case. Normalised mutual information provides consistently good estimates close to the

marker registration. There is one image pair (6) for which the estimate is significantly worse, and in this case the estimate for MR variance may in fact be closer to the marker estimate. Visual inspection of this dataset shows significant artefact in the PET image. In addition, visual comparison of the registration estimates indicates that the marker based estimate may be visually misaligned (see arrows on Figure 6.22). In this case it is not clear which estimate is actually superior.

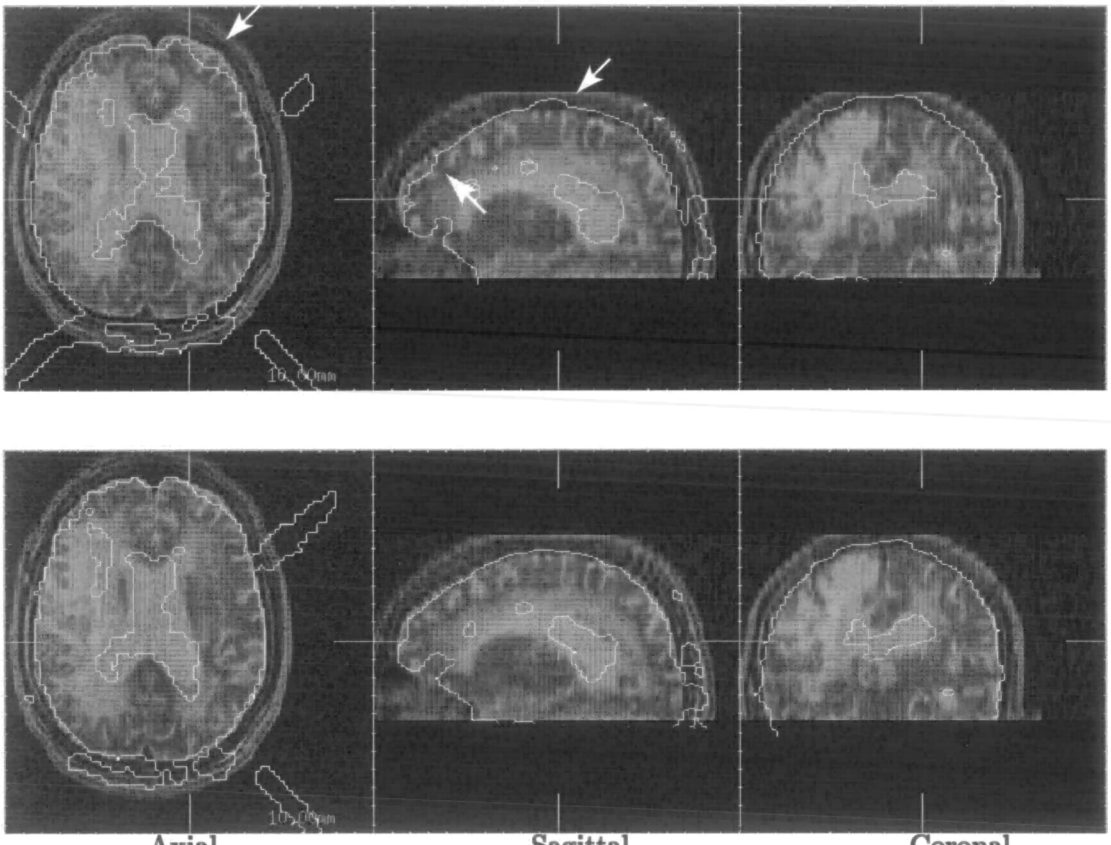


Figure 6.22: Slices through MR image volume for patient 6 illustrating registration estimate provided by optimisation of Normalised Mutual information (bottom) and the fiducial marker estimate(top).

6.6 Capture Range and Imaged Field of View

The experiments of the previous section compared the behaviour of the measures across a database of typical anatomical and physiological examples. The two other main variables determining the successful recovery of registration are the distance of the starting estimate away from true alignment, and the field of view (both axial and transaxial) of the image data. Varying the field of view of an image for a given registration estimate will exclude or include regions of material from the registration evaluation. This in turn will change the balance of values in the joint probability distribution and so will change which image values have most affect on the alignment measures. Targeted CT (and sometimes MR) images including only the inferior or superior regions of the skull are commonly acquired. These regions contain different amounts of shared structure capable of constraining the registration estimate.

This section will try to examine the behaviour of the measures when we vary the field of view of the images. This is achieved by a set of experiments where the quality of the starting estimate and field of view of the images is parameterised and varied. The performance of the measures is then compared by the number of successfully recovered alignments.

6.6.1 Exploration of Alignment Parameter Space

Rigid alignment between a pair of images requires the optimisation of 6 parameters. The set of experiments described in this section are essentially intended to compare the structure of this six dimensional parameter space for different measures and different image pairs. An ideal approach would be to evaluate this parameter space directly to compare the behaviour of the measures. Practical computational considerations mean that this is not possible, and we must find alternative approaches.

We would like to know at what size of initial misalignment an automated registration will fail. One approach to looking at this is to take an image pair with a good transformation estimate, randomly perturb this estimate by increasingly large amounts and record which of the starting estimates is recovered. This allows comparison of the parameter spaces provided by different measures with respect to the optimisation used, and so is a test of the system as a whole rather than simply the measure. This still provides useful information since, in the end, it is the behaviour of the complete system we are interested in.

Given the dimensionality of the parameter space and the computational cost in evaluating points in it, we must take care in the approach used to choose our starting estimates. It is impossible to adequately sample the space to get a complete idea of its

structure, but by considering what we want to know there are approaches to simplifying the problem. There are two fundamentally different types of information we might want to gain about the parameter space in these experiments.

Firstly, we are interested in the robustness of a system in recovering a large range of typical clinical misalignments (large scale structure of the parameter space), and secondly we are interested in the final precision of the alignment estimates (small scale structure). So, we would like to evaluate the performance of the system from a range of starting estimates at different distances (in parameter space) from the correct solution. We can therefore partition up the starting estimate space by the distance of an estimate from the correct solution.

If we treat all the parameters as equivalent we may simply use estimates that have a given displacement R (in parameter space) from the correct solution. This means that we simply want to choose points on a 6D hyper-sphere of radius R. Picking points on a 6 parameter sphere can produce estimates with very large translational components but small rotational components or large rotational and small translational components. In these set of experiments we have chosen to select rotational and translational components separately. Effectively this means that if we are picking random misalignments that are large, then it ensures both the rotational and translational components are large.

In order to create a set of translations of a given size in a random direction, we need to effectively pick points randomly on the surface of a 'translation' sphere. To do this we can pick two random angles, one ω between 0 and 2π and the other ψ between 0 and π. There is then the problem of the mapping of uniform random numbers onto the surface of this sphere. The mapping changes their uniform distribution and results in an uneven covering of the surface (particularly at the poles). To avoid this problem the angles ω and ψ may be picked with a pre-defined distribution:

$$\omega = \cos^{-1}(1 - 2v_1) \tag{6.3}$$

and,

$$\psi = 2\pi v_2 \tag{6.4}$$

Where v_1 and v_2 are random numbers between 0 and 1. These can then be used to locate 3D coordinates on a sphere of radius R:

$$\begin{aligned} t_x &= R\cos(\omega)\sin(\psi) \\ t_y &= R\sin(\omega)\sin(\psi) \\ t_z &= R\cos(\psi) \end{aligned} \tag{6.5}$$

where R determines the size of the translation. An equivalent approach can be used in

choosing a random rotation,

$$\begin{aligned} \theta_x &= R\cos(\omega)\sin(\psi) \\ \theta_y &= R\sin(\omega)\sin(\psi) \\ \theta_z &= R\cos(\psi) \end{aligned} \tag{6.6}$$

6.6.2 Experimental Design

The problem of recovering alignment is a function of both starting estimate and field of view of the image. In this experiment a large volume image pair was selected from the Vanderbilt Retrospective Registration Project database (the MR T1 and CT images for patient 5) with comparable extents in each axis. This provides an accurate, independent estimate against which successful and failed registrations can be judged. Sets of randomised starting estimates were then produced with increasing sizes of rotations and translations. In these experiments sets of 50 randomised transformations on 'translation spheres' of 10, 20, and 30mm and 'rotation spheres' of 10, 20 and 30 degrees were used.

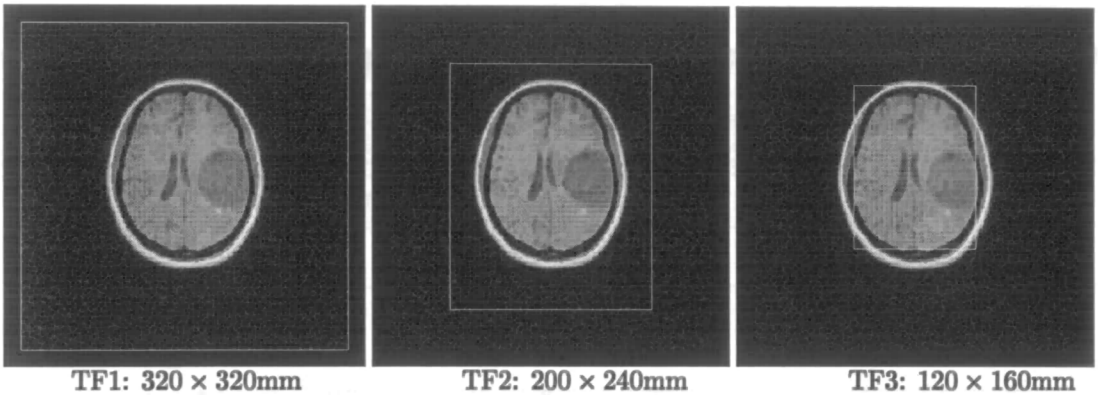
Transaxial Field of View

The transaxial (in-plane) field of view can vary considerably between clinical acquisitions. Images are acquired with a large field of view to include, for example, the structure of a stereotactic frame. For other cases a small field of view may be reconstructed to enlarge an area of clinical interest in the brain. The marker based registration image data all had a very large reconstruction field, compared to typical clinical images acquired at our site. In this experiment we will look at the affect of changing in-plane extent of the same image pair on the behaviour of different registration measures.

Varying the in-plane field of view will determine the proportion of the image containing air signal in the two modalities. In order to look at the effect this has on the registration measures, the images for patient 5 described in the previous section were truncated in-plane at two levels to give three pairs of images to register. The fields of view selected (labelled TF1, TF2 and TF3) in the MR, CT and PET images are shown in Figures 6.23, 6.24 and 6.25.

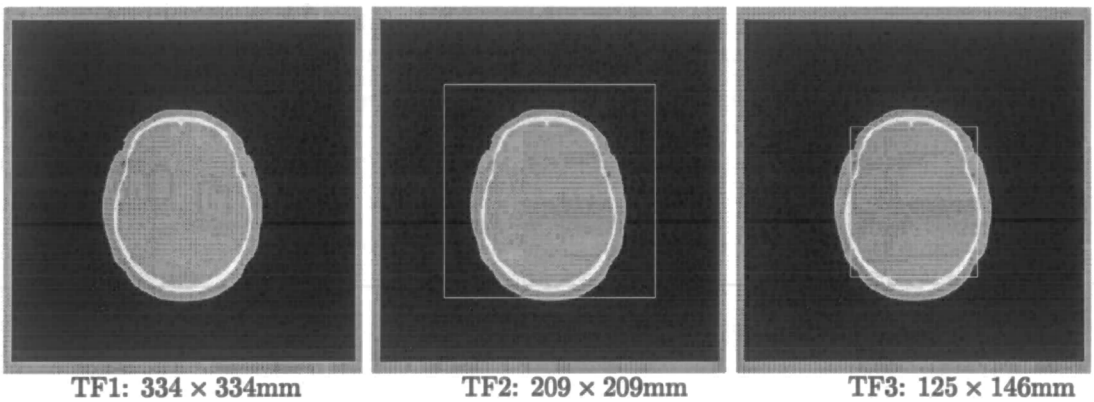
Axial Field of View

Axial targeting of MR-CT scans of the brain, particularly of CT, scans is very common. In this set of experiments the behaviour of the system for MR-CT registration using different axial extents will be examined. For clinical MR-PET studies this type of truncation is much less common, since the extent of the PET scan is usually fixed (because of the design of the scanner) to include the majority of the brain. In addition the MR images acquired for comparison with PET at our site are typically full brain



TF1: 320 × 320mm TF2: 200 × 240mm TF3: 120 × 160mm

Figure 6.23: The three levels of MR Transaxial field of view used for experiments.



TF1: 334 × 334mm TF2: 209 × 209mm TF3: 125 × 146mm

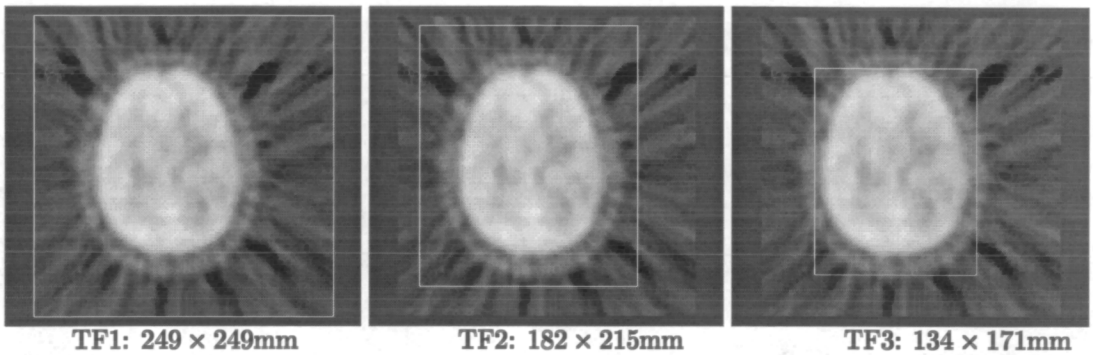
Figure 6.24: The three levels of CT Transaxial field of view used for experiments.

3D acquisitions. As a result this type of truncation for MR-PET registration was not investigated in this study.

The large volume MR-CT image pair used in the previous transaxial experiments was again used in these experiments. The image set with extent TF2 was truncated axially into three $4.4cm$ sections referred to as AFU, AFM and AFL (upper, middle and lower) which are illustrated in Figure 6.26.

6.6.3 Analysis of Results

In looking for the successful recovery of alignment we are not interested in differentiating those estimates which are close to registration, we simply need a way of classifying and counting the 'outliers' which are failures. If, following optimisation to some minimum step size, the estimate is still far from true alignment, we can assume that the estimate will not reach an acceptable solution and count it as a failure. In order to save computational costs in these experiments involving robustness, we therefore need not optimise the alignment using smaller and smaller steps to gain a very precise registration. In these experiments the step size was reduced to a minimum of only $\frac{1.50}{2^3}mm$ in translation and



TF1: 249 × 249mm TF2: 182 × 215mm TF3: 134 × 171mm

Figure 6.25: The three levels of PET Transaxial field of view used for experiments.

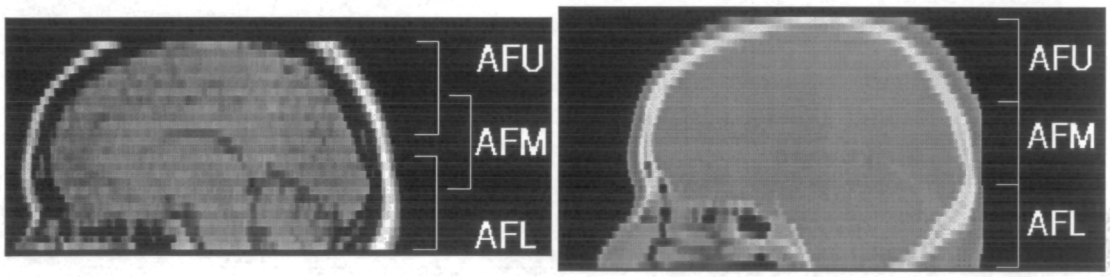


Figure 6.26: The three levels of axial field of view used for experiments.

$\frac{1.43}{2^3}°$ in rotation, using the same approach as described in section 5.6.2.

In order to then distinguish failures, the mean voxel displacement (see section 6.4.1) from the position defined by the marker estimate was evaluated and thresholded at a distance equal to the MR slice thickness ($4mm$). Examination of the distribution of estimates for the measures confirmed that this simple approach provided an effective differentiation between 'failures' and 'successes'.

6.6.4 Results: MR-CT

Transaxial Field of View

The number of successfully recovered misalignments, for each measure and each level of misalignment are shown in table 6.1. These show an appreciable difference in the behaviour of the system when using different measures. As expected from the single start estimates over the whole database, optimisation of MR variance provides a poor measure of alignment for all combinations of field of view and starting estimate. Only around half the closest starting estimates ($10mm$ and 10°) are recovered and a negligible proportion of those from further away.

Correlation coefficient of CT soft tissue with MR provides robust recovery of alignment for the two larger fields of view (TF1,TF2), but recovers very few of the starting

estimates for the smallest field of view (TF3). This indicates that this measure may be dependent on the skin/air boundary for the registration estimate which is mostly excluded by the smallest field of view. Minimisation of corresponding CT variance provides robust recovery for all starts, for all but the smallest field. Here with TF3, few of the two largest initial misregistrations are recovered.

Comparison of the three entropy measures with respect to truncation is interesting. Minimisation of joint entropy behaves similarly to the soft tissue correlation measure, providing robust recovery for all images except for the smallest field of view (TF3). Maximisation of mutual information conversely provides good recovery of alignment for all but the largest field of view (TF1). This confirms the simulation results of section 4.4 which indicated poor behaviour as the field of view is increased. The normalised mutual information measure $Y(M;N)$ provides overall the most robust recovery of alignment with only one failure for the largest initial misalignment and smallest field of view.

Successful Alignments with Varying Transaxial F.O.V.

Measure	10mm and 10deg			20mm and 20deg			30mm and 30deg		
	TF1	TF2	TF3	TF1	TF2	TF3	TF1	TF2	TF3
$\Gamma(M,C)$	50	50	3	50	50	8	50	50	4
$\sigma(C)$	50	50	50	50	50	21	50	50	8
$\sigma(M)$	24	29	33	0	1	8	0	0	2
$H(M,C)$	50	50	0	50	50	0	44	50	0
$I(M;C)$	42	50	50	16	50	50	3	50	48
$Y(M;C)$	50	50	50	50	50	50	50	50	49

Table 6.1: Number of successfully recovered MR-CT alignments (as defined in the text) from 50 random starts at three levels of misalignment, as transaxial field of view is varied.

Axial Field of View

The number of successfully recovered misalignments, for each measure and each level of misalignment are included in table 6.2. As with the transaxial fields of view, these show an appreciable difference in the behaviour of the system when using different measures.

Correlation of soft tissue values provides poor recovery of alignment of the central region of the head (AFM) for all starting estimates. The failed estimates indicate that the measure does not constrain alignment along the z axis (up-down). This may again be due to inadequate scalp-air boundary perpendicular to the z axis in the AFM volume. Minimisation of corresponding MR variance provides poor recovery of the smallest ($10mm$ and 10°) misalignments and fails for all the larger scale misalignments. Minimisation of joint entropy also provides poor recovery of truncated images, particularly for the two larger misalignments.

For the smallest misalignments ($10mm$ and 10°) both mutual information ($I(M;N)$ and normalised mutual information ($Y(M;N)$) provide good recovery of alignment of all the axial regions. As the misalignment increases the measures exhibit increasing numbers of failures, with significantly worse behaviour for the upper (AFU) and lower (AFL) regions.

Successful Alignments with Varying Location of Axial F.O.V.

Measure	10mm and 10deg			20mm and 20deg			30mm and 30deg		
	AFL	AFM	AFU	AFL	AFM	AFU	AFL	AFM	AFU
$\Gamma(M,C)$	45	0	47	25	1	16	7	0	6
$\sigma(C)$	45	46	39	18	32	18	0	7	3
$\sigma(M)$	2	31	15	0	0	0	0	0	0
$H(M,C)$	21	19	27	8	7	10	0	0	6
$I(M;C)$	46	50	30	20	48	13	12	39	8
$Y(M;C)$	41	50	50	18	45	43	6	22	28

Table 6.2: Number of successful MR-CT alignments (as defined in the text) from 50 random starts as the location of the imaged axial field of view is varied.

6.6.5 Results: MR-PET

Transaxial Field of View

The results of the MR-PET transaxial experiments are included in table 6.3. As with the experiments over the full database of images, the correlation coefficient measure has proved very poor and recovered no alignments at any of the fields of view or starting estimates. The variance measure $\sigma(M)$ provides good estimates for a large proportion of the two closer starting estimates at the larger fields of view. At the smallest fields of view and poorest starting estimates registration failures increase to a significant level. The opposing measure $\sigma(N)$ provides appreciably fewer acceptable estimates, recovering few of the poorer starting estimates.

Of the entropy measures, joint entropy only recovers a significant number of alignments at the closer starting estimates. Interestingly, the middle field of view (TF2) is very much worse than the larger (TF1) and smaller (TF2) fields of view at the $10mm$ or ° starting estimates. Mutual information provides very good recovery at the two smaller fields of view. As with the MR-CT case, the estimates from the larger field of view (TF1) confirm the normalisation problem identified in section 4.4.

Again, normalised mutual information provides the best overall behaviour with only 4 failures from the poorest starting estimates at the smallest field of view.

Successful Alignments with Varying Transaxial F.O.V.

Measure	10mm and 10deg			20mm and 20deg			30mm and 30deg		
	TF1	TF2	TF3	TF1	TF2	TF3	TF1	TF2	TF3
$\Gamma(M,N)$	0	0	0	0	0	0	0	0	0
$\sigma(M)$	47	50	50	38	49	40	36	50	24
$\sigma(N)$	28	19	25	24	14	10	15	14	0
$H(M,N)$	43	3	30	34	3	1	15	3	0
$I(M;N)$	27	50	50	21	50	50	8	50	49
$Y(M;N)$	50	50	50	50	50	50	50	49	47

Table 6.3: Number of successfully recovered MR-PET alignments (as defined in the text) from 50 random starts at three levels of misalignment, as transaxial field of view is varied.

6.7 Summary

The results of these experiments are the key in distinguishing the behaviour of the different measures of alignment. As we change the balance of materials in the images by modifying both the underlying anatomy and physiology (section 6.5), and the extent of the images (section 6.6), so the joint probability distribution of image values will change. This change then influences the response of the different measures to misalignment of shared structure in the images.

For MR-CT alignment, over the database of images, $\sigma(C)$ and $Y(M;C)$ provide comparable measures of alignment which are consistently better than those recovered using the other measures. Varying the axial and transaxial field of view indicates that normalised mutual information $Y(M;C)$ provides a significantly more robust measure of alignment. Over the 18 image pairs this measure provided a maximum MIPVD of $2.5mm$ and a minimum MIPVD of $0.4mm$ from the fiducial marker based estimate.

For MR-PET alignment $\sigma(M)$ and $Y(M;C)$ both provide good estimates of alignment over all the image pairs. As with MR-CT, when the transaxial field of view of the images is reduced, so the normalised mutual information is seen to provide a much more robust response. In this case over the 14 image pairs this measure provided a maximum MIPVD of $6.5mm$ and a minimum MIPVD of $1.4mm$ from the fiducial marker based estimate.

At smaller fields of view where air makes up a lower proportion of the image, mutual information and normalised mutual information both provide a very robust measure of misalignment. This confirms earlier findings on a database of non-stereotactic images with smaller fields of view [85, 86]. As the field of view is increased the problem of normalisation highlighted by the simulation in section (4.4) leads to a significant number of failures using mutual information.

Chapter 7

Retrospective Alignment of Clinical MR and CT of the Head

7.1 Introduction

In the previous chapter we identified that normalised mutual information provides, overall the most robust measure of alignment and provides comparable precision and accuracy to the other measures tested. In this chapter we are going to examine in more detail the application of this approach to the automated alignment of clinical MR and CT of the head. Essentially this chapter is concerned with the behaviour of the registration algorithm with respect to clinical MR-CT image characteristics and artefacts. The first section extends the work of the previous chapter by looking at the affect of image resolution and noise on the accuracy of the registration estimate with respect to the Vanderbilt maker based gold standard. These experiments, like those of chapter 5 make use of a controlled set of image data for which an independent estimate of registration is available.

The Vanderbilt data is very different from many of the images acquired clinically at our site, in particular in terms of:

- Image resolution
- Field of view
- Size and type of misalignment

The following sections of this chapter go on to look at the behaviour of the registration algorithm when applied to the retrospective registration of a larger range of clinical image data. The results in this chapter illustrate the experiences I have had in carrying out MR-CT alignment for a number of different surgery planning and guidance projects at UMDS. The task is one of providing visually the most accurate registration on data

containing a range of geometric and intensity artefacts. It depends heavily on the use of visualisation tools such as *rview* (See section 6.4.2) to provide a sensitive assessment of image alignment.

Section 7.3 includes a detailed description of clinical MR-CT data collected for this experimental evaluation and the range of artefacts which may influence the registration estimate. The registration algorithm is applied to recover the alignment of the 26 image pairs making up the database. Although we have no accurate independent ground truth estimate of alignment for this data, the automated estimates are compared to manual landmark based registration estimates. Visual inspection of estimates also provides a valuable indication of the performance of the system. A simple multi-start experiment to assess the precision of the parameter values calculated for every clinical image pair is also carried out. Although there are examples of images in the database which contain a number of specific geometrical artefacts which it is possible to address individually, here we are interested in providing a clinically useful registration with minimal user interaction.

One of the main differences between the registration of clinical data and the data used in the previous chapter is that the orthogonal scaling, and sometimes skew, parameters are known only with limited accuracy (or are known with great accuracy but are communicated with low accuracy). To achieve an accurate registration it became obvious that in some cases we would like to correct or at least check the validity of these parameters as part of the registration process. The registration algorithm was therefore extended to include the recovery of these parameters along with the conventional rigid body transformation. This extension is described in section 7.5. A comparison of estimates using only a rigid transformation to those including scaling and skew parameters is then included in section 7.6.

7.2 The Effect of Image Noise and Resolution on Alignment Accuracy

In this set of experiments we are going to investigate how the image noise and sampling resolution affects the accuracy of the estimate provided by optimisation of normalised mutual information. To do this we are going to take one of the images from the Vanderbilt study (The MR-T1 image pair of patient 5) with an accurate independent estimate of alignment, and change the noise level and spatial resolution.

7.2.1 Experimental Method

Image Sampling Resolution

Because of the multi-resolution approach to evaluation and optimisation of the measures, it is relatively simple to examine this relationship between image resolution and alignment accuracy. We can simply vary the finest cubic sampled image resolution used in the optimisation scheme (see chapter 4). In this experiment an image registration was carried out with finest cubic voxel resolutions of between $4.0mm$ (equivalent to the slice thickness) and $0.75mm$ at steps of $0.25mm$. Using the same sampling scheme as described in chapter 4, linear interpolation is used to increase resolution and Gaussian blurring is used to reduce resolution from the original image sampling rate. Example slices through images at two resolutions are shown in Figure 7.1.

In order to look at both the accuracy of a final estimate and its precision with respect to the marker estimate, 16 random starts were used at each resolution. For each of these estimates the MIPVD (see section 6.4.1) was evaluated with respect to the marker estimate to give an overall measure of registration accuracy.

Image Noise

In order to investigate the effects of measurement noise on the precision of the registration estimates, a range of Gaussian noise signals were added to both the MR and CT values at the original image resolution. Although this is not an accurate model for the MR noise, which requires the addition of Gaussian noise to the complex representation of the image, it still provides a valuable indication of the response of the system to errors in intensity values. The Gaussian distributions used were with mean values of zero and a range of standard deviations set to varying percentages of the range values in each image, from 0% to 100% in steps of 10%. Examples of the resulting images are shown for two levels in Figure 7.2.

A simple estimate of the levels of noise present in the original images was carried out by measuring the distribution of values in a uniform region in the images. The range of

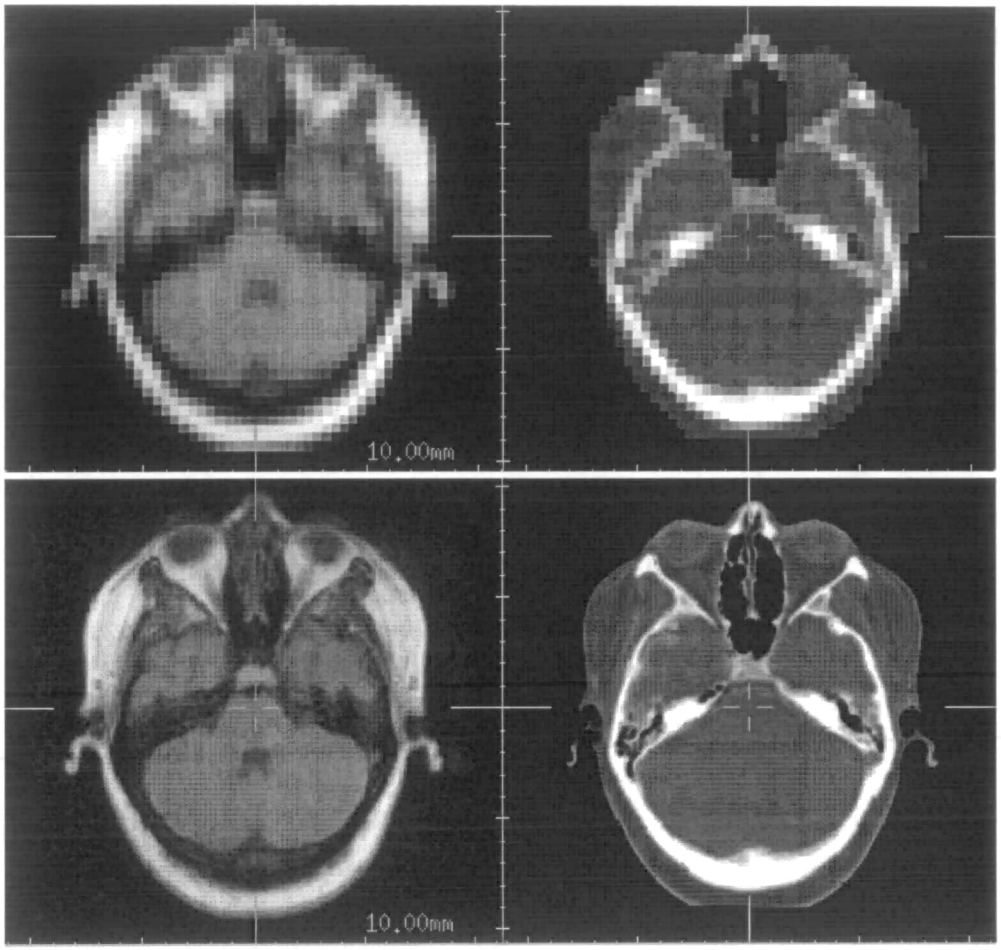


Figure 7.1: Examples of two image resolutions used to investigate registration accuracy, left MR , right CT. Upper image is with $4mm$ cubic voxels and lower image is with $1mm$ cubic voxels.

values in the original MR was 1746 and the values recorded in a region of air in the image had a standard deviation of around 10 units (0.57% of the range of values). The range of values in the original CT image was 3064 Hounsfield units and the distribution of noise values measured in the air had a standard deviation of 4.8 Hounsfield units (0.16% of the range of values). The levels of noise added are therefore considerably higher than that already present in the data.

In order to do the registration, a maximum cubic sampling rate of $1.5mm$ was used. As with image resolution, in order to look at both the accuracy of a final estimate and its precision with respect to the marker estimate, 16 random starts were used at each resolution, and the MIPVD evaluated for each estimate.

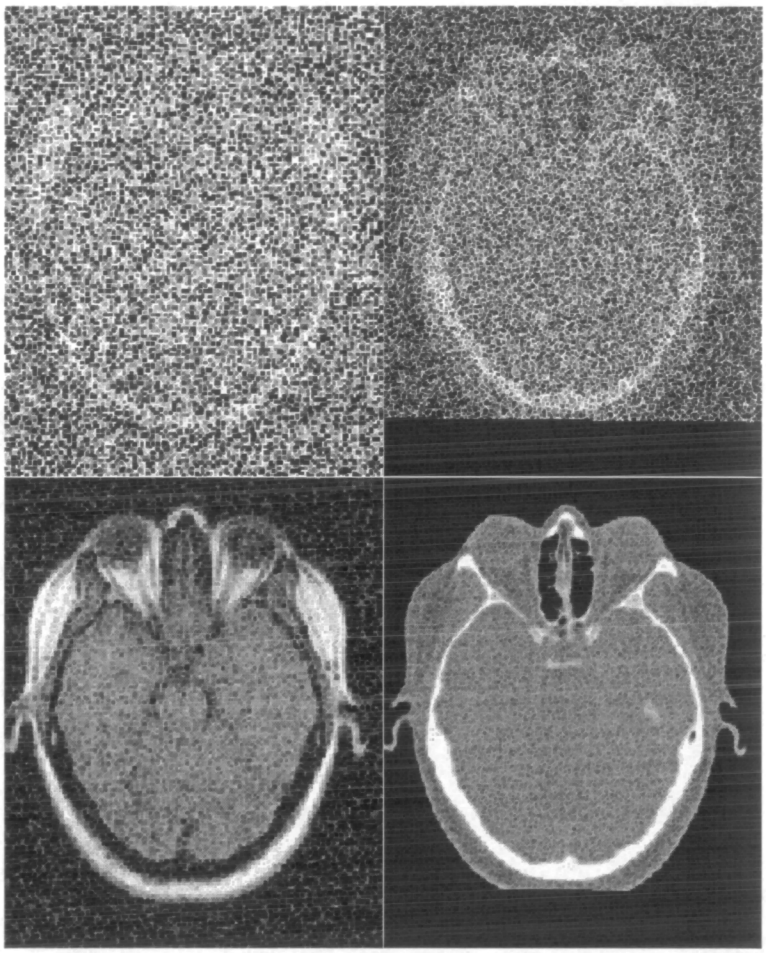


Figure 7.2: Examples of two levels of additive Gaussian noise used in experiments, left MR , right CT. Upper image is with added Gaussian noise with a standard deviation of 80% of the dynamic range. Lower image is with added Gaussian noise with a standard deviation of 10% of the dynamic range.

7.2.2 Results: Image Resolution

The errors in each estimate from 16 random starts at each image resolution are summarised by the graph of the averaged error for these starts in Figure 8.3. This shows an overall fall in voxel misalignment as the sampling resolution of the images is increased, but the relationship between alignment accuracy and voxel size is not smooth or monotonic. This may well be attributable to the small number of starts used to investigate the registration precision, but may also be due to the blurring out of corresponding boundaries at different resolutions in the two modalities. At the highest image resolution ($0.75mm$) the MIPVD averaged over 16 starts between the automated and marker estimates is $0.26mm$, with a standard deviation of $0.024mm$. This is within the expected accuracy of the fiducial marker estimate.

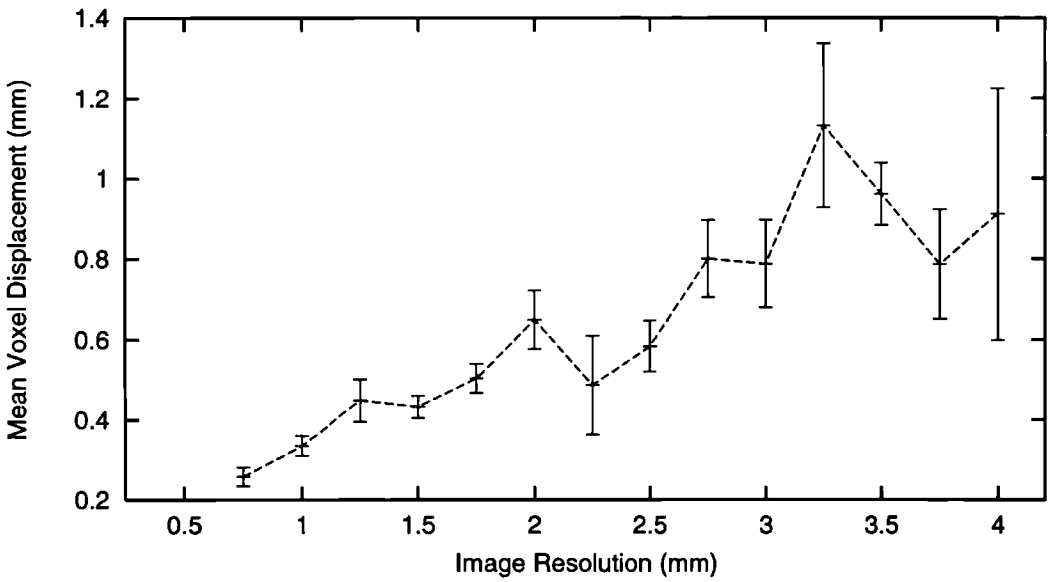


Figure 7.3: MR(T1)-CT Vanderbilt patient 5: MIPVD between the estimate provided by normalised mutual information and the marker estimate as cubic sampling resolution is varied: Average with standard deviation (error bars) over 16 starts at each sampling resolution.

7.2.3 Results: Measurement Noise

The errors in each estimate from 16 random starts at each noise level are summarised by the graph in Figure 7.4. There is an overall decline in both the accuracy and precision of the registration estimate as the added noise level is increased. Given the relatively high levels of noise used in this experiment the precision and accuracy of the final estimate remain quite high. Up until around a noise signal with a standard deviation of half the dynamic range, the displacement error is within around $1.5mm$ from the marker estimate. Such a noise level would mean the images would be of little clinical use for typical diagnostic imaging applications.

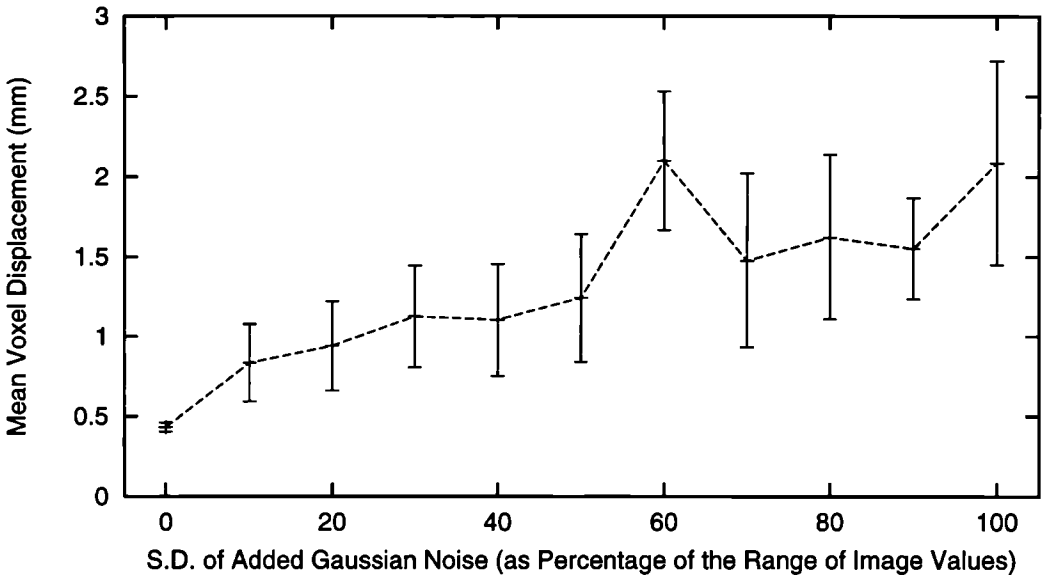


Figure 7.4: MR(T1)-CT Vanderbilt patient 5: MIPVD (at 1.5mm cubic image resolution) between estimate provided by normalised mutual information and the marker estimate as noise level is varied: Average with Standard Deviation (error bars) over 16 starts at each noise level.

7.2.4 Discussion

These two experiments have investigated the response of the system to image noise and resolution with otherwise ideal image data. In general, the level of noise acceptable in clinical imaging will not appreciably influence the registration estimate. Even at high levels of noise though, the registration accuracy and precision may be acceptable for some specialised applications. An example may be the alignment of intra-operatively acquired low-dose CT or fast acquisition interventional MR, with conventional pre-operatively acquired images.

The precision with respect to image resolution though is important for two reasons.

Firstly the inherent resolution of the clinical images may determine the limits of the registration accuracy. Secondly, the resolution at which we evaluate the similarity measure is important in the setting up of an efficient multi-resolution optimisation scheme. During the multi-resolution optimisation we can use this information to decide at what step size to terminate optimisation, making best use of the information at a given image resolution.

The highest image resolution used in the optimisation will determine the final registration accuracy. In practice, it is the computational expense of evaluating a similarity measure at high image resolution which is currently the limiting factor, and not the inherent resolution of the clinical MR-CT images.

Although resolution and noise impose fundamental limits on the registration accuracy, as we shall see in the following experiments, in clinical use the actual registration accuracy will often be determined by other, less predictable variables.

7.3 A Database of Clinical MR-CT Images

Any approach to defining an estimate must impose some constraints on the imaging protocol to enable an estimate of registration to be derived. The aim of the experiments in the rest of the chapter is to investigate the behaviour of the alignment algorithm using normalised mutual information when used to retrospectively align clinical image data. This image data is again all of the head but, it has been acquired for a number of mostly non-stereotactic procedures. In many cases the data was not specifically acquired for registration, and so enables us to look at the problems of true retrospective image registration over a range of clinical imaging protocols. The data can be split into two main groups which will be described in more detail in the following subsections.

7.3.1 Data Acquired Locally

The first group of data used for the tests consists of sixteen pairs (A to P) of clinically acquired MR and CT images of the brain. The images for patient A and G were acquired for the planning of epilepsy surgery. Image pair G includes a number of brain implants visible in both MR and CT. Patient H was imaged to plan the treatment of a carotid aneurysm. Patient I was suffering from a significant mycotic sinus infection visible as enhancing regions filling large areas of the air sinuses. The images were used in the planning of removal and treatment of the infection. The remainder of the images were acquired for the planning of the surgical removal of skull base tumours including an acoustic neuroma and a number of meningiomas of varying size.

The MR images consisted of:

- T1 weighted Gadolinium enhanced 2D multi-slice acquisitions (B,C,D,E,H,I)
- T1 weighted 3D gradient echo images (A,F,G,K-O)
- The background signal of a 3D phase contrast angiogram[1] (J)
- T2 weighted 2D acquisition (P).

All images had nominally transaxial slices except for the images of patient F which were nominally coronal.

The CT datasets range in size from the largest (patient A) to the more common targeted volumes of patients B to J (see tables 7.1 and 7.2). The acquisition of whole brain CT scans is extremely rare at our site. The CT volume for patient A was synthesised

[1]It is a conventional approach at our site to make use of the MR signal component of MR angiography (which is acquired simultaneously with the flow image and is therefore in registration with it) to enable the registration of MR flow images with CT or other MR.

by combining 11 slices acquired around the skull base and 17 slices acquired around the top of the head (both with 512×512 pixel $3mm$ slices), with seven 256×256 $6mm$ slices (which were tri-linear interpolated to 512×512 $3mm$ slices) acquired through the centre of the skull. As with the MR, CT slicing was nominally transaxial except for patient F which was coronal.

MR Image Volume Parameters

Patient	F.O.V. (mm)	Voxel size (mm)	Scan
A	240×220×186	0.938×0.938×1.5	axial
B	235×235×108	0.918×0.918×3.0	axial
C	210×210×90	0.820×0.820×3.0	axial
D	230×230×90	0.898×0.898×3.0	axial
E	220×220×21	0.859×0.859×3.0	axial
F	220×220×248	0.859×0.859×2.0	coronal
G	220×220×186	0.859×0.859×1.5	axial
H	220×220×72	0.859×0.859×0.9	axial
I	200×200×120	0.780×0.780×2.0	axial
J	240×240×124	0.938×0.938×1.0	axial
K	220×220×180	0.859×0.859×1.2	axial
L	220×220×156	0.859×0.859×1.2	axial
M	220×220×186	0.859×0.859×1.5	axial
N	220×220×168	0.859×0.859×1.2	axial
O	220×220×186	0.859×0.859×1.5	axial
P	180×180×57	0.352×0.352×2.5	axial
Q	200×200×64	0.781×0.781×1.0	axial
R	200×200×192	0.781×0.781×1.5	axial
S	220×220×150	0.859×0.859×1.5	axial
T	220×220×94.8	0.859×0.859×1.2	axial
U	230×230×180	0.898×0.898×1.2	axial
V	200×200×80	0.781×0.781×0.8	axial
W	220×220×144	0.859×0.859×1.2	axial
X	250×250×192	0.977×0.977×1.2	axial
Y	250×250×192	0.977×0.977×1.2	axial
Z	250×250×192	0.977×0.977×1.2	axial

Table 7.1: MR volume parameters.

7.3.2 Image Data from the COVIRA Database

The second group of image data originates from the COVIRA project database and consists of another 10 image pairs. The COVIRA Project was partly funded by the European Community and partly by industrial collaborators. The image data has in general a higher spatial resolution and field of view, particularly for the CT data, than the locally acquired image data. Tables 7.1 and 7.2 detail the image volume parameters. The later 3 image volumes (X,Y,Z) were acquired with a stereotactic frame and both MR

CT Image Volume Parameters

Patient	F.O.V. (mm)	Voxel size (mm)	Scan
A	250×250×126*	0.488×0.488×3.0	axial
B	250×250×84	0.488×0.488×3.0	axial
C	250×250×75	0.488×0.488×3.0	axial
D	250×250×57	0.488×0.488×3.0	axial
E	250×250×27	0.488×0.488×1.5	axial
F	250×250×102	0.488×0.488×3.0	coronal
G	250×250×55.5	0.488×0.488×1.5	axial
H	250×250×39	0.488×0.488×1.5	axial
I	250×250×60	0.488×0.488×3.0	axial
J	250×250×49.5	0.488×0.488×1.5	axial
K	250×250×31	0.976×0.976×1.5	axial
L	250×250×134	0.488×0.488×2.0	axial
M	250×250×75	0.488×0.488×3.0	axial
N	200×200×130	0.391×0.391×2.0	axial
O	220×220×102	0.488×0.488×3.0	axial
P	180×180×43	0.351×0.351×1.0	axial
Q	227×227×67.5	0.710×0.710×1.5	axial
R	208×208×87	0.406×0.406×1.0	axial
S	220×220×73.5	0.688×0.688×1.5	axial
T	200×200×87	0.625×0.625×1.0	axial
U	210×210×99	0.410×0.410×1.5	axial
V	160×160×74	0.312×0.312×1.0	axial
W	210×210×100	0.410×0.410×1.5	axial
X	300×300×147	0.586×0.586×1.5	axial
Y	300×300×147	0.586×0.586×1.5	axial
Z	300×300×147	0.586×0.586×1.5	axial

Table 7.2: CT image volume parameters, *after interpolation as described in the text.

and CT have a large field of view and high resolution. The MR acquisitions are all T1 weighted Gadolinium enhanced images, often containing significant regions of enhancing skull base lesion. The majority are 3D acquisitions. The CT scans are in some cases post-operative and include changes (in particular bone structures such as the skull) with respect to the MR acquisition.

7.3.3 A Summary of Clinical Image Characteristics

In applying a registration algorithm to a range of clinical image data we are looking at its behaviour with respect to a significant number of acquisition related variables in addition to variations in patient anatomy and image alignment. These are independent of the actual structures being imaged inside the patient and are a function of imaging protocol and clinical background.

Image Extent

The main variable in clinical axial MR-CT acquisitions of the brain, which has a large affect on the registration problem, is the axial extent of the image. This can be particularly limited in CT where X-ray dose is to be minimised. The range of axial extents of the image data used in these experiments is illustrated in Figure 7.5.

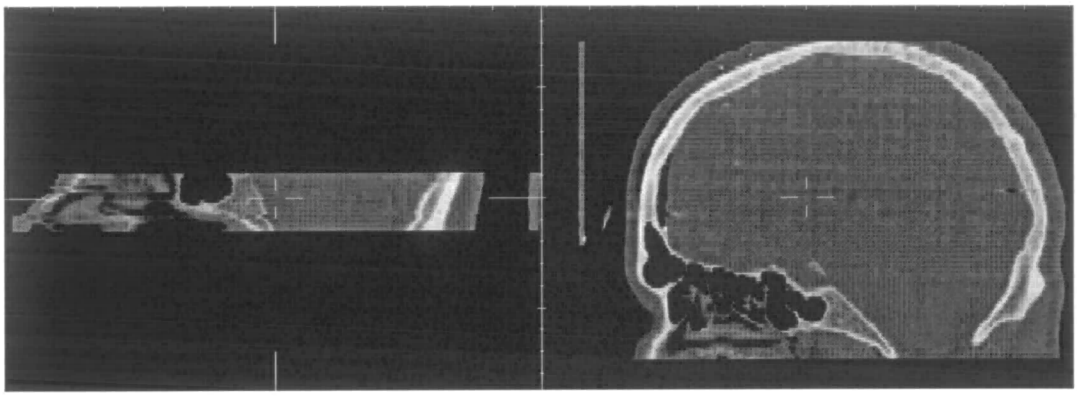


Figure 7.5: The axial extent of CT brain images varies considerably from full head stereotactic images (W,right) to targeted acquisitions (E, left).

Another related variable is the extent of the in-plane CT reconstruction. Image set V has a limited field of CT reconstruction which excludes an appreciable portion of skin surface as shown in Figure 7.6. Section 5.3.2 details the approach used to distinguishing image padding areas, and to resampling these non-rectangular regions.

Metallic Components

One important artifact results from the presence of metal components in the imaged field of MR and CT which can occur when the patient is imaged with a stereotactic frame. This causes significant localised geometric distortion in MR and intensity based reconstruction artefact in CT. The clinical data used in these experiments include a number of examples of this, one such is illustrated in Figure 7.7. Another source of artefact is the presence of metal in the fillings of teeth, This can be seen in the CT coronal slices of patient F in Figure 7.8. This is particularly a problem for coronal type slices where the in-plane reconstruction artefact may extend up into the skull and brain.

Tissue Deformation

Non-rigid 'distortion' between the two scans can also be caused by the soft-tissue deformation. This is always a problem for the skin surface of the scalp, particularly at the

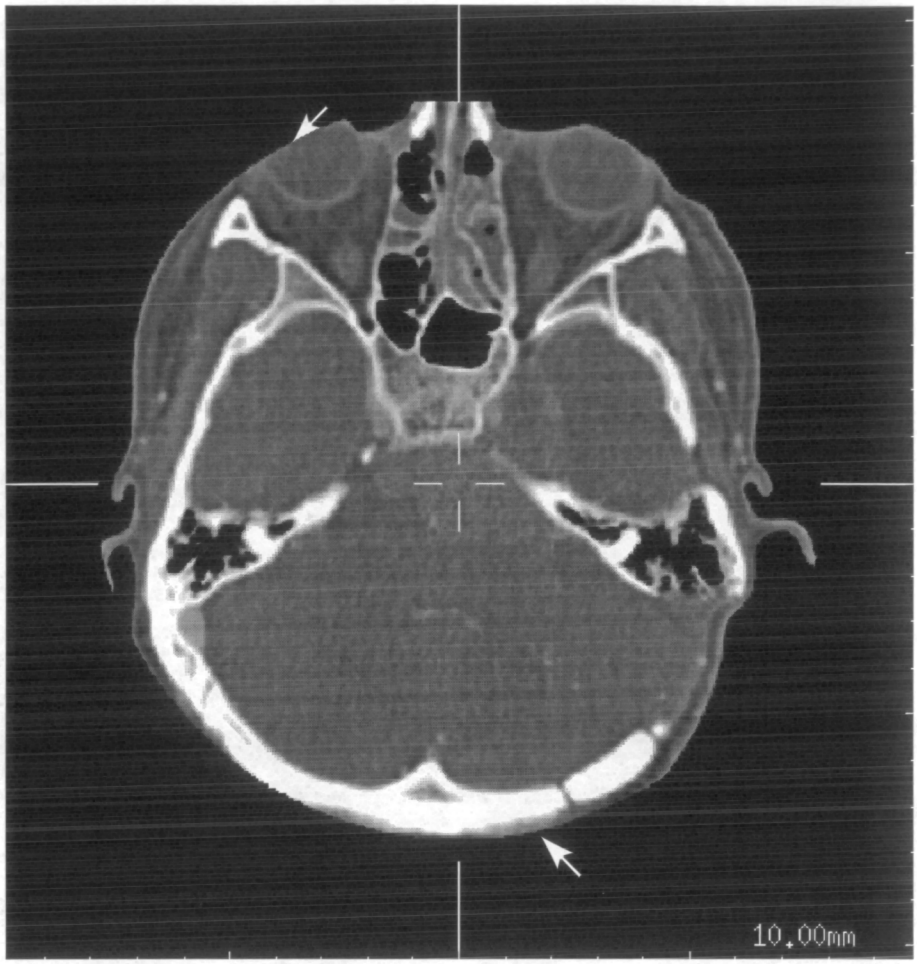


Figure 7.6: Example slice through image set V showing limited in-plane extent of CT image reconstruction with respect to skin surface.

back of the neck. The problem becomes more severe when both scans include regions lower in the head, such as the mouth, lower jaw and neck. An example of such an acquisition is image pair N which is illustrated in Figure 7.9. Here we are interested again in the registration of the skull base but structures in the neck have been included to aid in planning the surgical approach. This is also a problem in the coronal scans of patient F, where the jaw is differently positioned in the MR and CT. It is possible to simply exclude these regions from the registration, but in these experiments we are interested in the overall behaviour of the registration algorithm without such modifications. The inclusion of regions in the neck is also interesting in terms of the similarity measures, since in the neck we have significantly different structures imaged to those normally found in the brain.

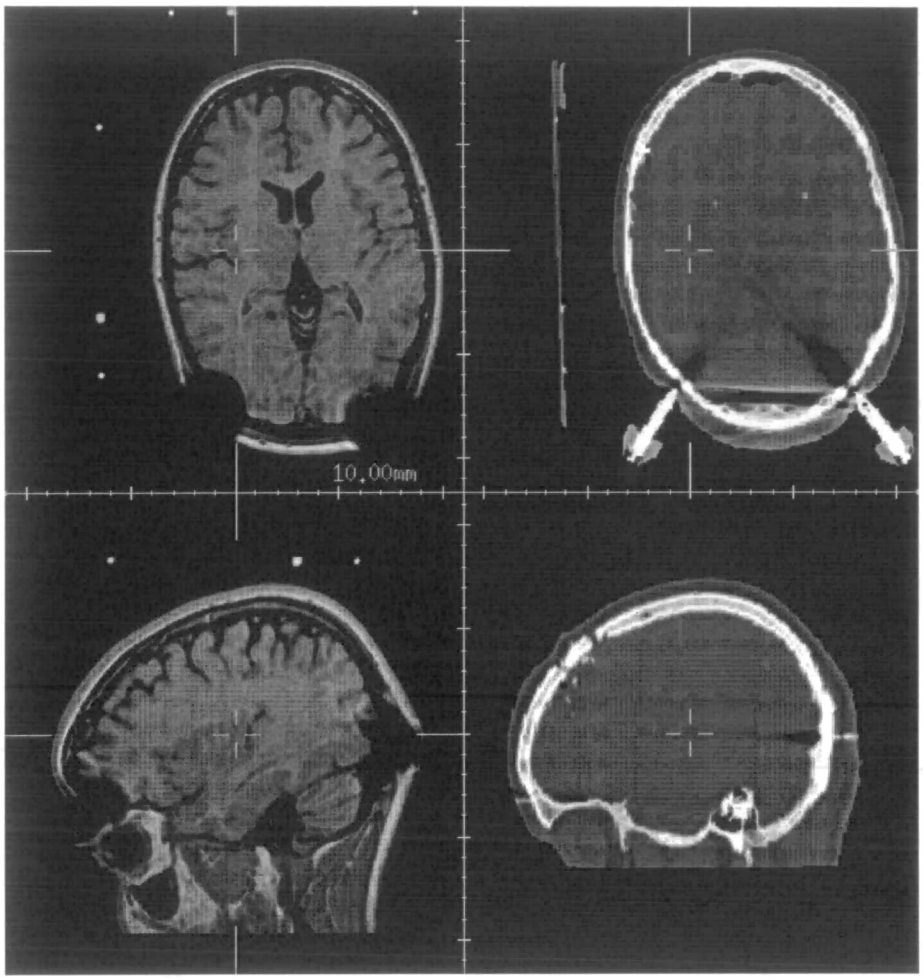


Figure 7.7: Example orthogonal slices through MR (left) and CT (right) images containing artefacts due to the presence of metal components in the stereotactic frame.

Patient Movement

Another important artefact occurring in clinical data results from patient movement during the image acquisition. It's specific form is dependent on both the nature of the patient motion during the acquisition, and on how the data is acquired over time. In CT where slices through the patient are generally acquired sequentially it is common to see a discontinuity between slices as shown for patient W in Figure 7.10.

In MR, because of the much larger range of acquisition modes, the nature of motion artefact is more variable. Some examples of artefact in the 2D slice-interleaved MR images used in these experiments are shown in Figure 7.11. For these types of artefact a global rigid transformation does not completely describe the spatial relationship between the two modalities. For cases such as that illustrated in Figure 7.10, a partitioning of the imaged space into multiple rigid transformations would provide a significant improvement. Such an approach though is even more difficult to validate. In these experiments

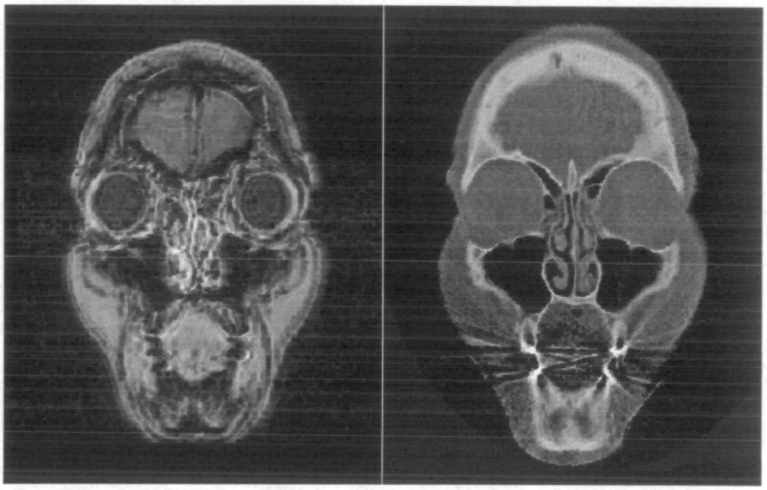


Figure 7.8: Example coronal slices through MR (left) and CT (right) images of patient F containing artefacts due to the presence of metal in the fillings of teeth.

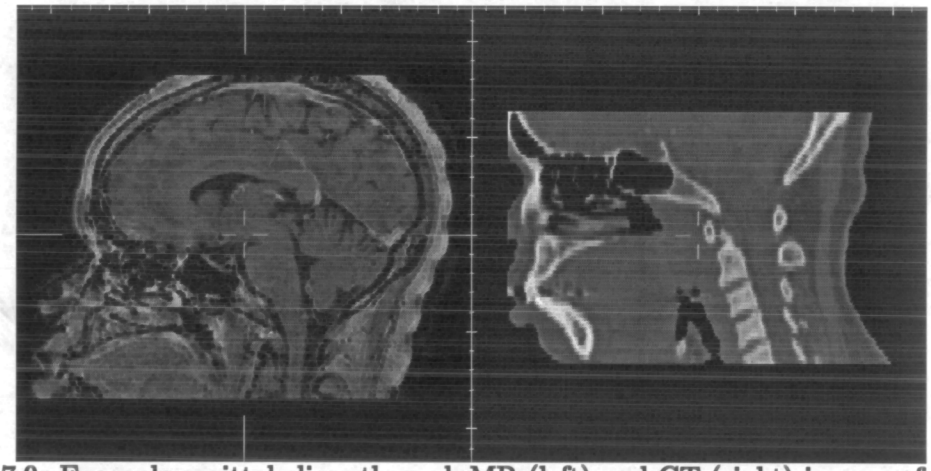


Figure 7.9: Example sagittal slices through MR (left) and CT (right) images of patient N showing acquisitions of the skull base including a significant portion of deformable structures in the neck and mouth.

we are simply seeking the best global rigid transformation providing a clinically useful accuracy, comparable to the manual point based registration currently in use.

7.3.4 Independent Manual Estimate of Registration

Group 1

Rigid registration estimates for the 16 locally acquired datasets (A-P) were obtained by identifying between 8 and 15 corresponding anatomical point landmarks [47]. These landmarks have been regularly used for rigid registration of skull base MR and CT at our site. It is important to note that the landmarks tend to be concentrated in the skull base, and so the registration will tend to be most accurate in this area. The points were

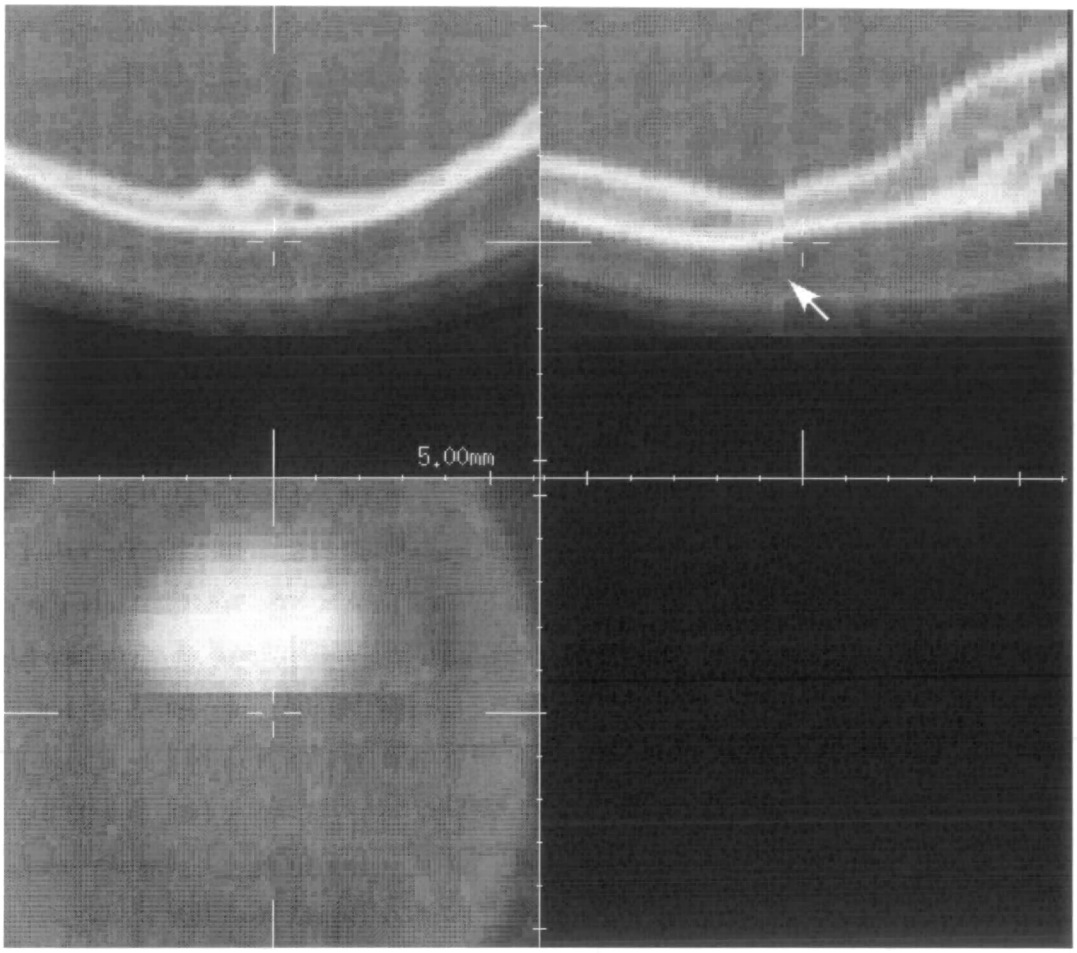


Figure 7.10: Example slices through the CT images of patient W showing movement during an acquisition.

aligned by minimising the distance between them using singular value decomposition. The resulting transformation parameters are shown in table 7.3. The root mean square (RMS) error values for these registration solutions are between 1mm and 6mm. If the point location errors are random, then the registration accuracy within the volume circumscribed by these points can be considerably smaller than this.

The registration solutions were visually inspected and considered to have similar accuracy to the point registration solutions that have been used on dozens of patients at our institution over several years.

Group 2

The second group of data was also registered interactively by corresponding point landmark identification, but as part of a study into the accuracy of a commercial image registration system. The study involved the registration of each dataset five times by four trained observers with a clinical background. The transformations used as a refer-

ence in the following experiments are the average of the four observers and five trials for each dataset.

7.3.5 Common Types of Clinical Image Misalignment

The manual results provide a valuable indication of the scale of typical misregistration encountered between brain images in clinical practice. Due to the nature of the acquisitions the image pairs have a range of misregistrations, some of which may be outside the capture range of the system. A good guide to the scale of misregistrations for each image pair is given by the manual estimates for groups 1 and 2 in table 7.3. The direction of the translation and rotation parameters used in this table and the rest of the thesis are illustrated between a reference MR image and a CT image in figure 7.12.

Translational misalignment can vary considerably and in the axial direction can be constrained by the axial extent of the images. One of the main features of non-stereotactic axial CT is that it is commonly acquired in a plane orientated to avoid the eyes, which are particularly sensitive to X-Ray exposure. In addition, the design of the head rest of CT scanners also contributes to a different orientation of the slicing plane. This is reflected in the θ_x (nodding) angle in table 7.3. The other two rotations are commonly much smaller since they are constrained in a similar way by the head rest in both MR and CT. The large rotational misalignments are a major difference from the Vanderbilt dataset used in the previous chapter, where the rotations were limited to one or two degrees in each axis.

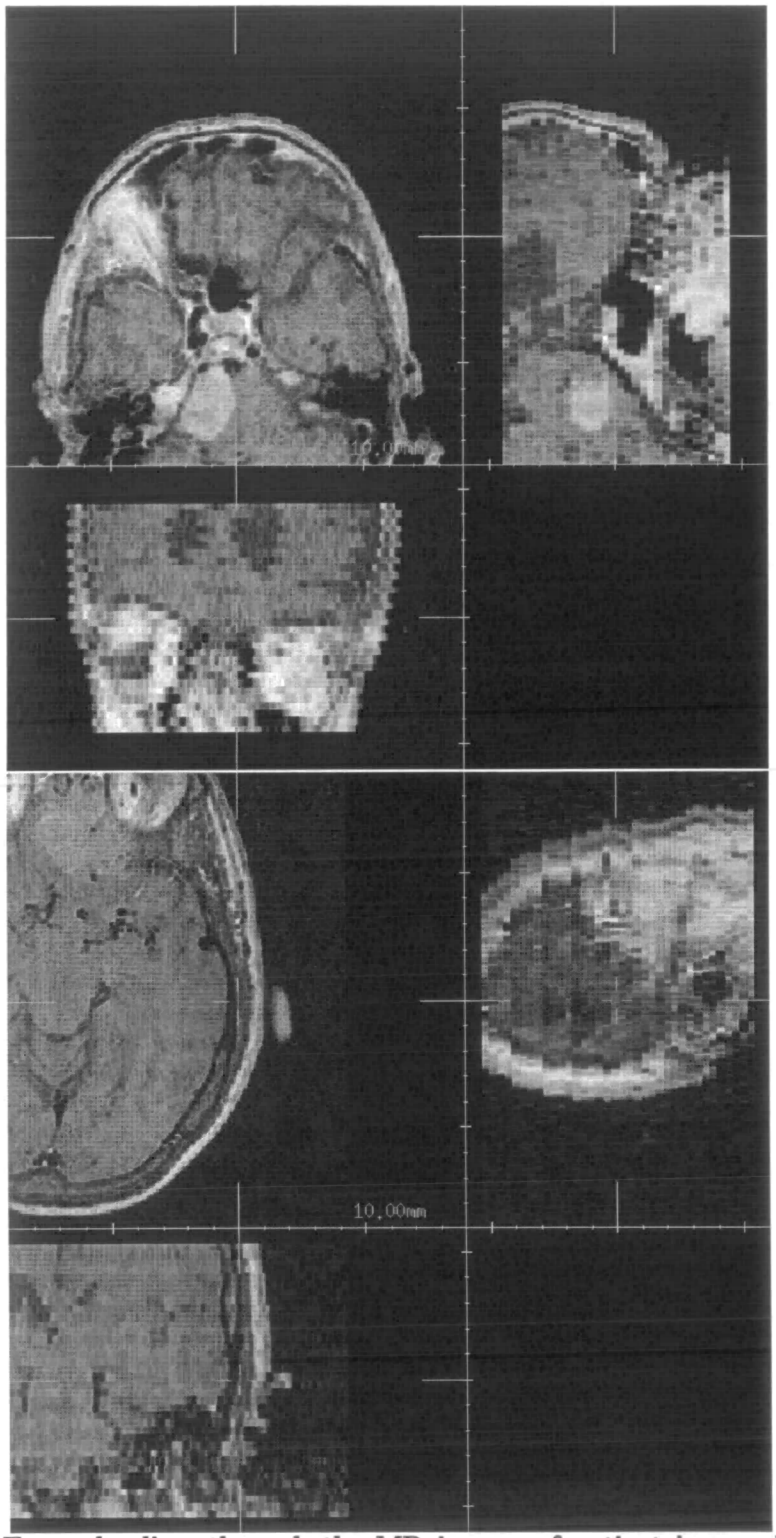


Figure 7.11: Example slices through the MR images of patient images B(lower) and D(upper) showing different artefact due to movement during 2D slice interleaved acquisitions.

Manual Registration Estimates

Patient	Translation (mm)			Rotation (deg)			R.M.S.	Num
	t_x	t_y	t_z	θ_x	θ_y	θ_z	Error	Points
A	8.38	33.68	2.15	24.06	-6.44	-0.87	2.56	12
B	3.56	23.97	-0.18	17.77	-0.59	-5.42	3.15	15
C	0.31	12.04	-9.49	-1.33	-0.72	2.33	1.57	14
D	-0.38	13.89	4.86	0.90	-4.41	4.85	2.37	11
E	5.83	-0.09	-3.30	11.41	1.15	0.89	2.58	10
F	6.43	7.62	38.55	-2.36	-2.58	3.42	5.51	11
G	0.89	19.69	-23.16	27.34	-2.90	-0.07	3.24	8
H	-4.18	11.65	5.19	4.36	-0.72	2.05	2.03	11
I	0.77	2.26	-22.98	2.12	-0.40	0.10	1.45	12
J	-2.05	14.30	-19.26	31.80	-0.47	1.16	2.12	12
K	-1.89	9.96	-44.45	35.45	-0.50	4.79	3.20	7
L	-3.06	-3.08	-23.83	2.48	-1.36	0.17	1.54	9
M	3.87	7.70	-22.18	25.43	-6.35	-2.71	1.79	8
N	2.88	2.68	-59.21	11.58	2.17	0.44	1.45	8
O	10.41	20.34	-19.91	25.13	0.56	2.50	1.77	9
P	0.36	2.58	-5.39	12.42	6.65	0.28	2.20	9
Q	4.23	5.31	7.68	13.82	1.75	4.03	–	12
R	-3.16	-16.30	-31.20	7.43	-2.17	-5.32		12
S	-2.79	0.04	-19.0	13.55	4.23	-3.59	–	12
T	1.12	-4.85	-25.60	7.57	0.83	0.04	–	12
U	-6.18	-14.50	-44.80	11.61	1.63	1.25	–	12
V	-0.12	10.27	-0.94	13.67	-3.74	0.86	–	12
W	5.29	-2.00	-23.10	9.27	-0.17	-6.97	–	12
X	1.03	1.66	9.95	-0.43	1.88	1.40	–	12
Y	-0.23	-4.59	-15.60	-0.13	0.68	-0.98	–	12
Z	5.91	-2.41	-6.60	-0.14	0.57	-2.07	–	12
Max	10.41	33.68	38.55	35.45	6.56	4.85		
Min	-6.18	-16.30	-59.21	-2.36	-6.44	-6.97		

Table 7.3: Manual registration estimates by identification of corresponding anatomical landmarks

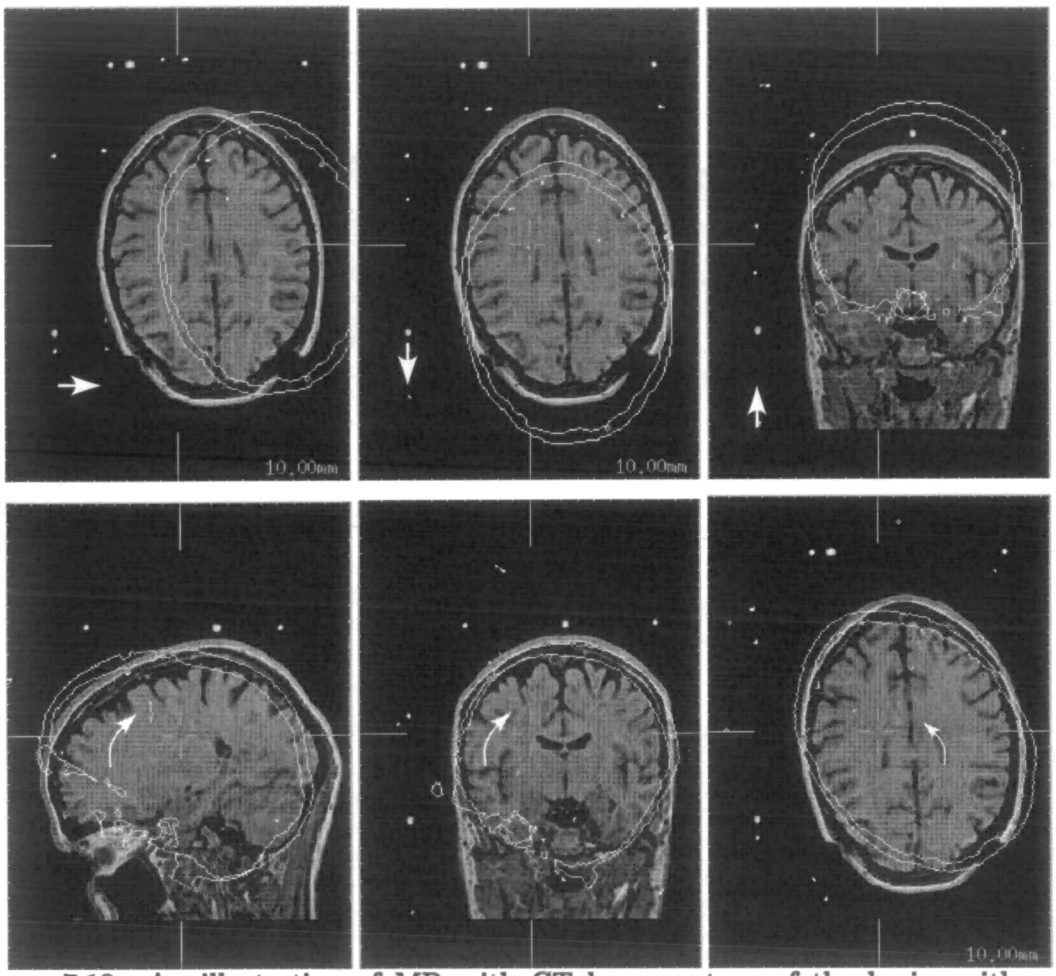


Figure 7.12: An illustration of MR with CT bone contour of the brain, with misalignments in different rigid parameters. Top from left to right: transaxial slice with $t_x = +30mm$, transaxial slice with $t_y = +30mm$, coronal Slice $t_z = +30mm$. Bottom from left to right: sagittal slice with $\theta_x = +30mm$, coronal slice with $\theta_y = +30mm$, transaxial slice with $\theta_z = +30mm$.

7.4 Automated Rigid Registration

7.4.1 Algorithm Parameters

In order to ensure a fair test of automated alignment, identical parameter settings for the registration algorithm were used to align all the image pairs. Optimisation was achieved by reducing the step sizes of the parameters as described in section 5.6.2. A graph of MIPVD in Figure 7.13 shows the errors with respect to the marker estimate for the Vanderbilt image data used in section 7.2.1, as the number of intensity bins representing values in the two modalities is varied (using $1.5mm^3$ minimum voxel size). This indicates that the choice of the number of bins used has little effect on the final registration accuracy. The number of bins used for these experiments was 64.

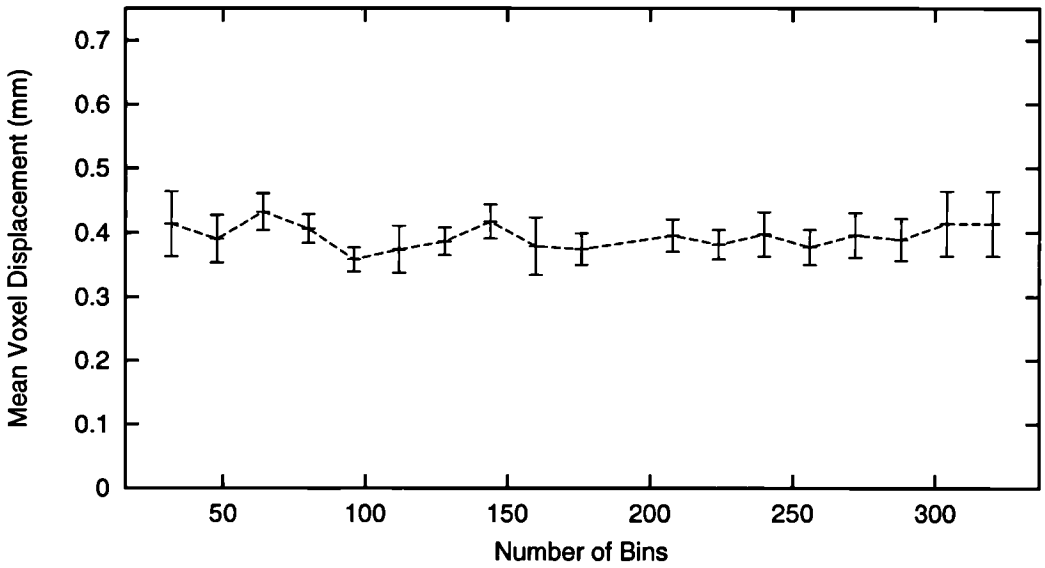


Figure 7.13: MIPVD between the estimate provided by normalised mutual information and the marker estimate as the number of intensity bins (for both images) in the joint probability distribution is varied: Average with standard deviation (error bars) over 16 starts at each number of bins.

7.4.2 Experimental Design

In this experiment we are interested in two things, firstly if the automated registration algorithm can recover a visually acceptable registration solution for each clinical image pair, and secondly an indication of how precise this estimate will be (we do not have a gold standard to determine true accuracy). In experiments in the previous chapter the parameter space was investigated by randomly perturbing an estimate and re-registering a large number of times. Here, an initial registration was carried out using 6 parameters from the default (no rotations and image centres aligned) starting estimate. This gives

an indication of the robustness of the system to the conventional starting estimate. This automated estimate was then randomised to form 16 starting estimates using random translations of $5mm$ and random rotations of 5° (randomised as described in section 6.6.1).

7.4.3 Accuracy

In all cases except patient E (the smallest axial extent), a visually acceptable registration was provided by automated registration. Figure 7.14 shows the MIPVD (calculated as described in section 6.4.1) between the automated and manual point based estimates for each image pair. The majority of the automated estimates fall within an MIPVD of $3mm$ of the point based estimate, with nine of the estimates being within $1mm$. Further

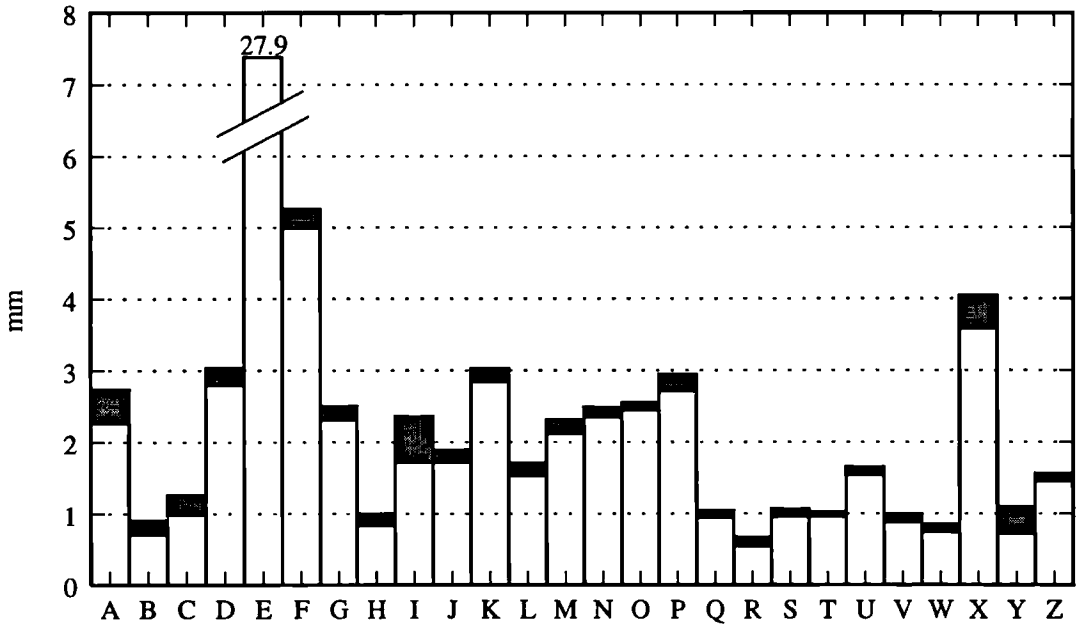


Figure 7.14: MIPVD in mm (truncated at 8mm) between rigid transformation estimates provided by manual point based alignment and automated optimisation of normalised mutual information from a single start (Standard deviation for each image pair from 16 starts around this is shown as error bars).

experimentation indicated that it was not possible to robustly recover alignment of image pair E because of the limited axial extent. Alignment was only possible with improved starting estimates and appreciably reduced search parameters.

There are a number of examples where visual inspection indicates that the rigid estimate provided by automated alignment is significantly superior to that provided by manual point based alignment. An example is patient G shown in Figure 7.15, where the MIPVD between the two estimates is $2.4mm$. The Figure shows that there is misalignment of the bone contour at the back of the skull base when overlayed on the

MR, as indicated by the white arrows. In this case manual landmark identification was made difficult because of the difference in orientation (the nodding angle: θ_x) of the slice planes. A second example is set I shown in Figure 7.16, where misalignment of the bone

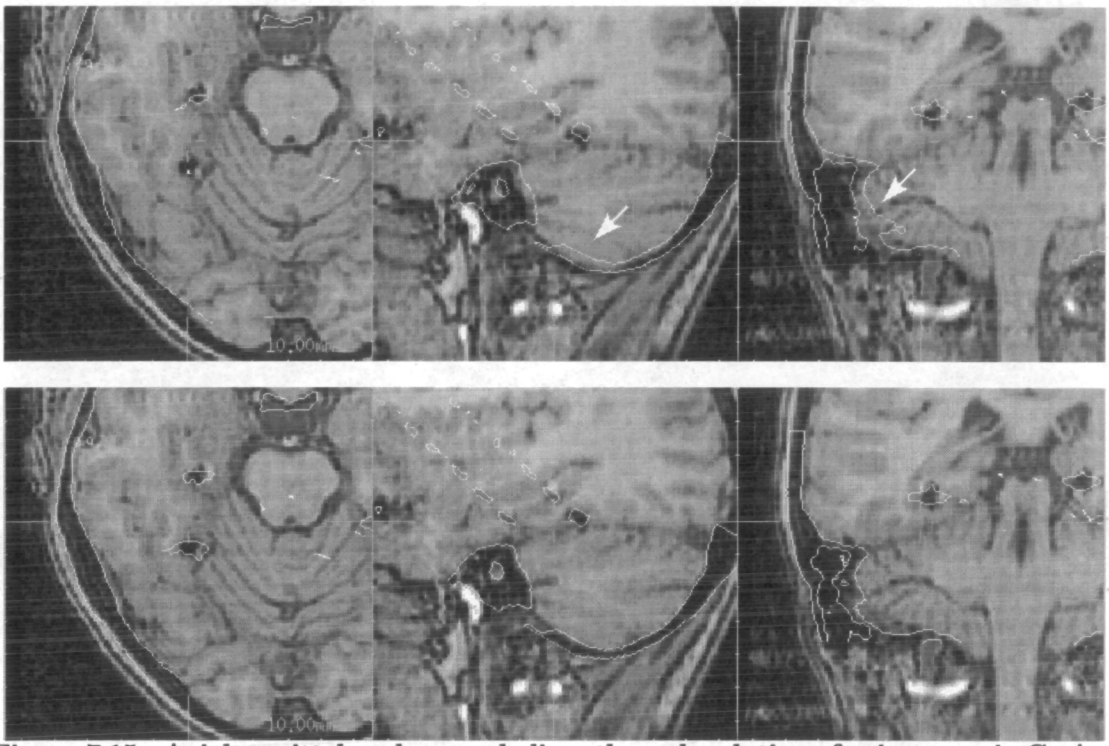


Figure 7.15: Axial, sagittal and coronal slices through solutions for image pair G: An example where the estimate provided by automated registration (bottom) is visually superior to that provided by manual point based registration (top). White arrows indicate regions of misaligned bone contour from CT.

of the eye socket can be seen in the manual estimate.

There are two pairs (F,X) which have automated estimates with a mean voxel displacement greater than 3mm from the point based estimate. Closer visual inspection of these showed that these image pairs may have significant errors in the scaling estimates supplied with the data. Because the manual estimates were provided by points selected in the skull base, and the automated estimate was derived from the whole data, there is a resulting difference in the registration estimates. This problem will be covered in more detail in the section 7.6.

7.4.4 Precision

The estimates of the precision of the transformation parameters for each of the image pairs is summarised by their standard deviations from 16 starts in Figure 7.17. As can be seen there is a significant variation in the spread of estimates for each of the image

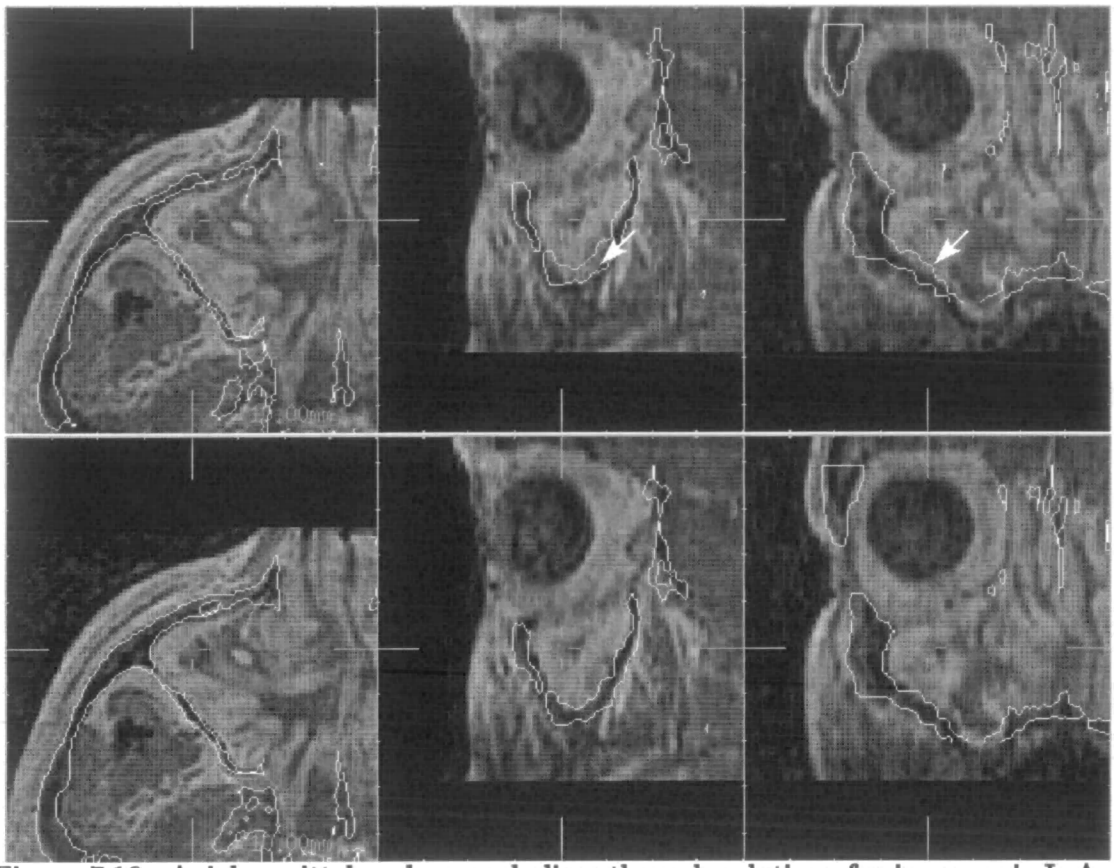


Figure 7.16: Axial, sagittal and coronal slices through solutions for image pair I: An example where the estimate provided by automated registration (bottom) is visually superior to that provided by manual point based registration (top). White arrows indicate regions of misaligned bone contour from CT.

pairs. The z axis translation estimates appear to have a larger spread than the in-plane translations. In addition the x and y rotation parameters (out of plane rotations) appear to be less well constrained. This may reflect the lower spatial resolution of the images in the between-plane axis. This affect though will also be influenced by the actual mis-alignment of the imaging planes of the two modalities.

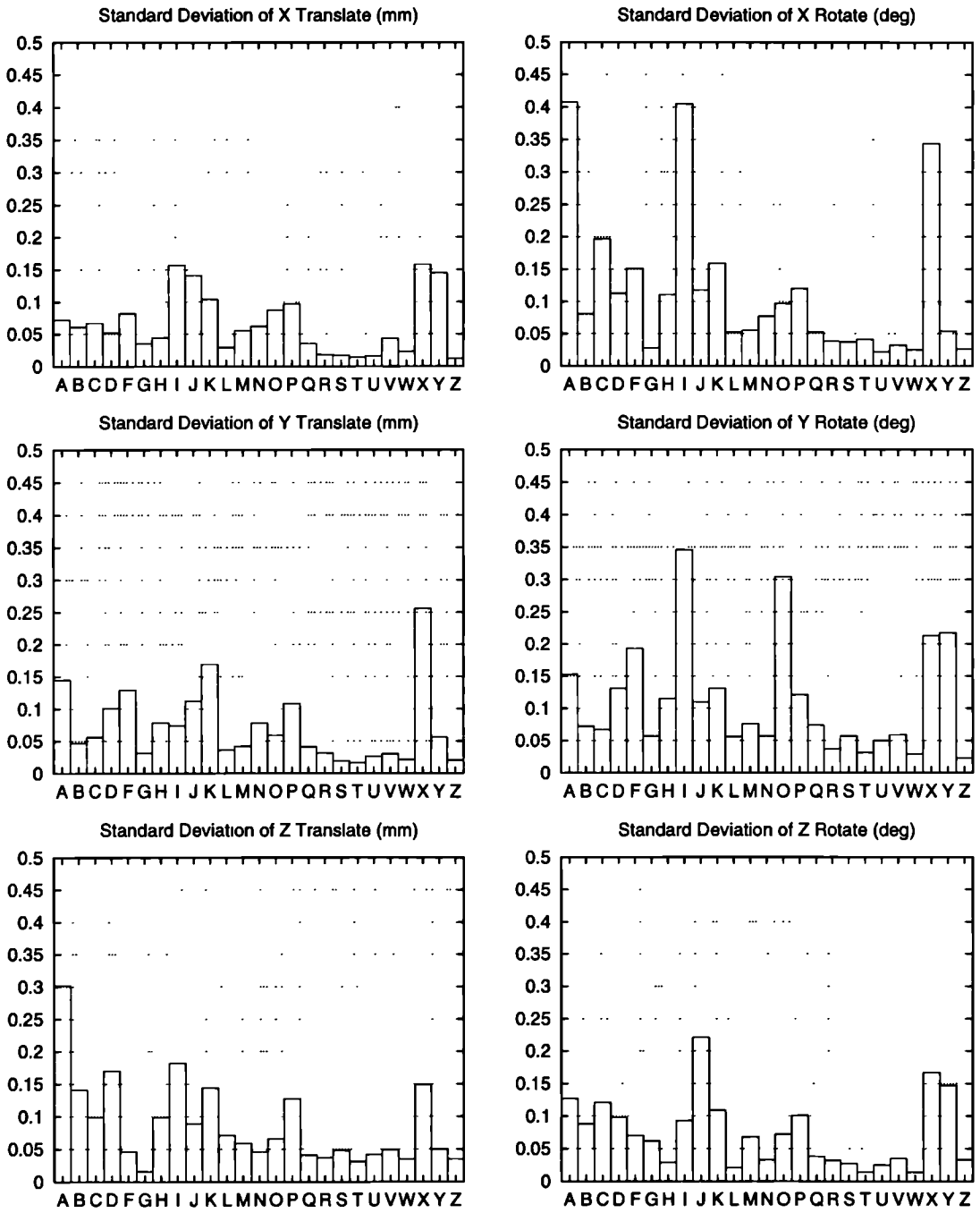


Figure 7.17: The standard deviation of rigid registration parameters estimated for each clinical image pair (except pair E).

7.5 Recovery of Scaling and Skew Parameters

Having developed tools to enable the accurate visual inspection of image registration results, it became apparent that in a number of cases the accuracy of a rigid alignment estimate between MR and CT data was significantly affected by the orthogonal scaling parameters of the two images and the skew (gantry tilt) of the CT volume.

7.5.1 Why Include Scaling and Skew Estimates?

In general, estimates of orthogonal scaling and CT gantry skew will be provided from the scanner settings and calibration, by the image header information (for example within the DICOM format). In some cases (for example when data has been imported by scanning films or in raw digital format from other hospitals) this information is not available. Even when scanner estimates of scaling and gantry tilt are available there are limits in the accuracy of those values. MR images in particular have scaling estimates accurate only to within a few percent leading to significant uncertainty in the distance between structures over larger fields of view. In such cases if the structures within the images constraining alignment are some distance from the structures of interest, then the resulting misalignment may have a significant affect on the clinical interpretation of registered images.

7.5.2 Incorporating Scaling and Skew in the Transformation Matrix

Both scaling and skew can be expressed in a 4×4 homogeneous transformation matrix form as used for the rigid transformation parameters. The orthogonal scaling $\{\gamma_x, \gamma_y, \gamma_z\}$ in each axis can be represented by a single matrix,

$$T_{SCALE}(\gamma_x, \gamma_y, \gamma_z) = \begin{bmatrix} \gamma_x & 0 & 0 & 0 \\ 0 & \gamma_y & 0 & 0 \\ 0 & 0 & \gamma_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The gantry tilt is simply a skew in the $Y - Z$ plane of angle ϕ_{zy}. The y location of a point in the image is proportional to its location along the z axis (the plane of the bed in CT). This can be expressed by a matrix:

$$T_{SKEW}(\phi_{yz}) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \tan\phi_{zy} & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The rigid registration algorithm was then extended to include scaling and gantry tilt by simply constructing these additional matrices from the extra parameters, and multiply-

ing the rigid transformation matrix T_{RIGID} by these,

$$T(t_x, t_y, t_z, \theta_x, \theta_y, \theta_z, \gamma_x, \gamma_y, \gamma_z, \phi_{yz}) = T_{RIGID}.T_{SCALE}(\gamma_x, \gamma_y, \gamma_z).T_{SKEW}(\phi_{yz}) \quad (7.1)$$

7.5.3 Optimisation of Scale and Skew Parameters

As with translations and rotations, the main question for optimisation is how to relate changes in each parameter. With scaling, we can evaluate the displacement of a point at some distance from the origin. If we take, as we did with rotations and translations, a point $x = 60mm$ from the origin, then to achieve a displacement of 1mm will require a change in scale factor of $\frac{100}{60}\%$. Similarly with skew angle, we can evaluate the displacement at a distance of 60mm,

$$\delta\phi = \tan^{-1}\left(\frac{1}{60}\right) = 0.954^0 \quad (7.2)$$

These are illustrated in Figure 7.18. This simple relationship between parameters was

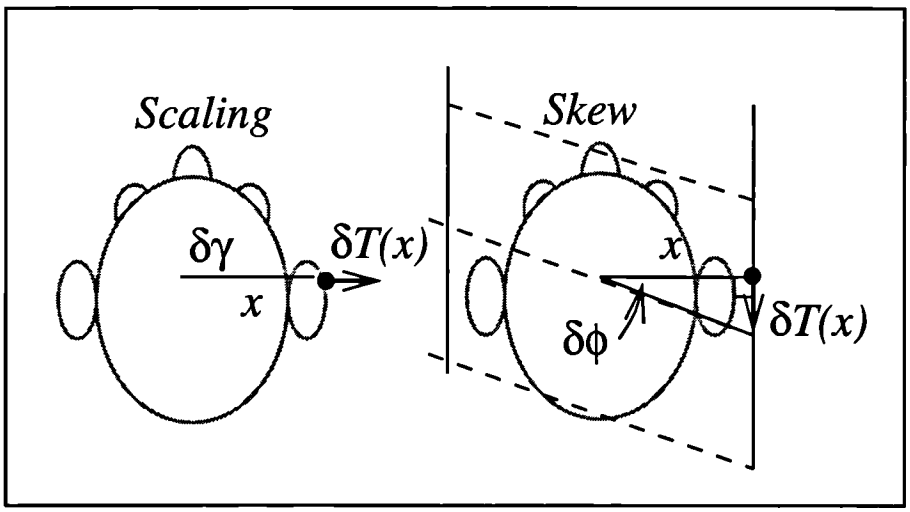


Figure 7.18: The affect of small changes in skew and scale on the displacement of a point at a given distance from the origin.

found to provide acceptable performance in optimisation. The specific values were varied over a small range (±25%) and found not to be critical to the outcome of the optimisation.

As with rigid registration, the parameters can be optimised using a multi-resolution approach, where parameter step size is related to image resolution (section 5.6.2). In this case though we can assume that the scale and skew estimates have only a few percent error and so limit optimisation of these extra parameters to small steps at the highest image resolution. This was found to significantly improve the robustness of the optimisation.

7.6 Scaling and Skew Recovery from Clinical Data

7.6.1 Experimental Design

As with the rigid estimates, we can carry out a registration using 10 parameters on each of the image pairs from the default starting estimate. This gives an indication of the robustness of the system for typical starting estimates over the full database of images.

From this estimate 16 randomised starting estimates were produced. Each starting estimate has scale and skew estimates derived from the first estimate. These are formed from the 16 combinations of change in scale and gantry tilt of $\delta\gamma_x = \pm 2\%, \delta\gamma_y = \pm 2\%, \delta\gamma_z = \pm 2\%, \delta\phi_{yz} = \pm 2°$. For each of these 16 estimates, the rigid translation and rotation parameters are perturbed from the values in the first registration by a random translation of $5mm$ and a random rotation of 5° (the same perturbations being used for each image pair). These 16 transformations were then used as starting estimates to the optimisation from which 16 independent estimates were recovered to give an indication of precision. These can be directly compared to the precision of the rigid estimates.

7.6.2 Results

Detection of Scaling and Skew Errors

The estimates for the scaling and gantry tilt parameters for each image pair are shown in Figure 7.19. The estimated gantry tilt angle is generally less than around 2° except for four or five datasets. For the images D,F and K the gantry angle of the CT had been lost or omitted in data transfer.

The in-plane (x and y) scaling parameters are small and within those expected due to MR distortion. The between plane (z) scaling 'error' estimates are generally much greater. One of the larger values (patient L) was attributed to poor calibration of the spiral CT bed speed. The speed was judged to have a 7% error by the service technician, which agreed well with the automated estimate of scaling. The images for patients D and F show a z (slice thickness) scaling error which was attributable to MR distortion.

The affect of the estimation of these extra parameters on the visually judged quality of the registrations varies. Figure 7.20 shows an example where the addition of extra parameters to the transformation estimate is globally visible because of the large axial extent of the data. It is interesting to look at the alignment of the stereotactic frame components for the 10 parameter solution for image pair X. This is illustrated by the slices in Figure 7.21. From these it is obvious that the visibly acceptable estimate for the patient structures does not apply as well to regions away from the centre of the dataset. The presence of a frame would remove the possibility of patient motion, indicating that

this misregistration is probably due to more complex geometric distortions. Although the alignment using scale and skew is visibly superior to the rigid estimates, further visual inspection shows that some misregistration is still present in the patient images, even far from the metal components of the frame.

More complex geometric distortion may well be the cause for misalignment in image pair T, illustrated in Figure 7.22. Here visual inspection indicates misalignment around the left eye socket for both the rigid and scale-skew estimates. The quality of the MR and CT scans indicate that patient motion may not have been a factor in this mis-alignment.

Precision

We can look at the precision in two ways, firstly the precision with which the extra parameters are estimated, and secondly what affect the extra parameters have on the precision of the rigid components of the transformation. The precision of the scale and skew parameters are summarised by the graphs of their standard deviation over 16 starts in Figure 7.23. From this it can be seen that the in-plane scaling parameters (X and Y) are quite precisely defined, with standard deviations of fractions of a percent in scale for most of the image pairs. The estimates for Z axis scale and gantry tilt overall are much more poorly defined, with standard deviations of two or more percent in scale and up to a degree in gantry tilt. The spread of estimates appears greater for those images with limited axial extent. The poor estimates may be attributed to the shape of structures being imaged, coupled with the limited extent of the data. In addition, the limited spatial resolution in the axial direction may also contribute to the fall in precision.

The precision of the rigid parameters in the 10 parameter registration is appreciably effected, with particularly the X-axis (nodding) rotation being very poorly defined in comparison to the 6 parameter rigid transformations. This rotation will be closely related to the gantry tilt and Z-axis scaling parameters.

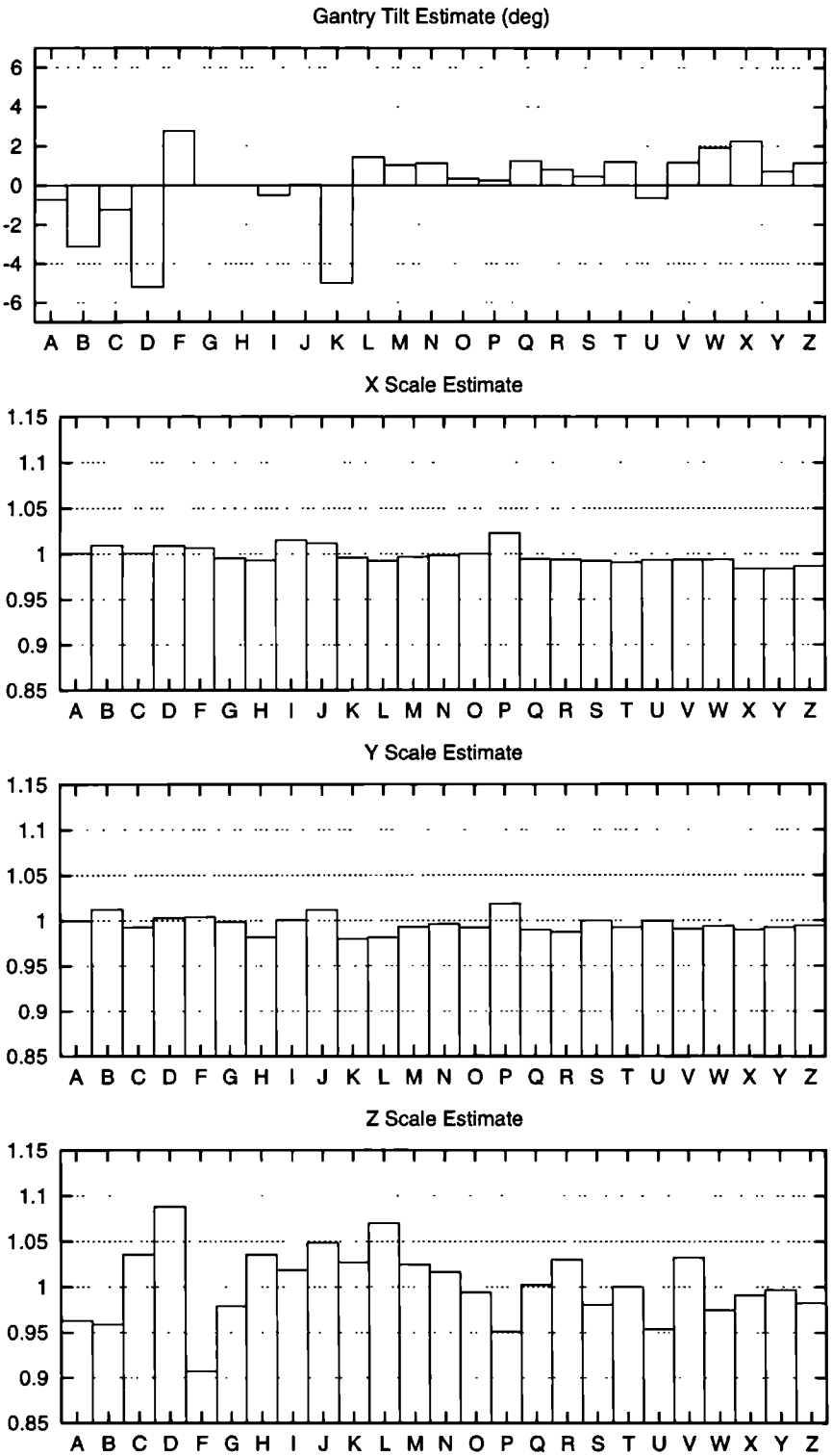


Figure 7.19: Orthogonal scaling and gantry tilt estimates for each image pair (from 16 starts).

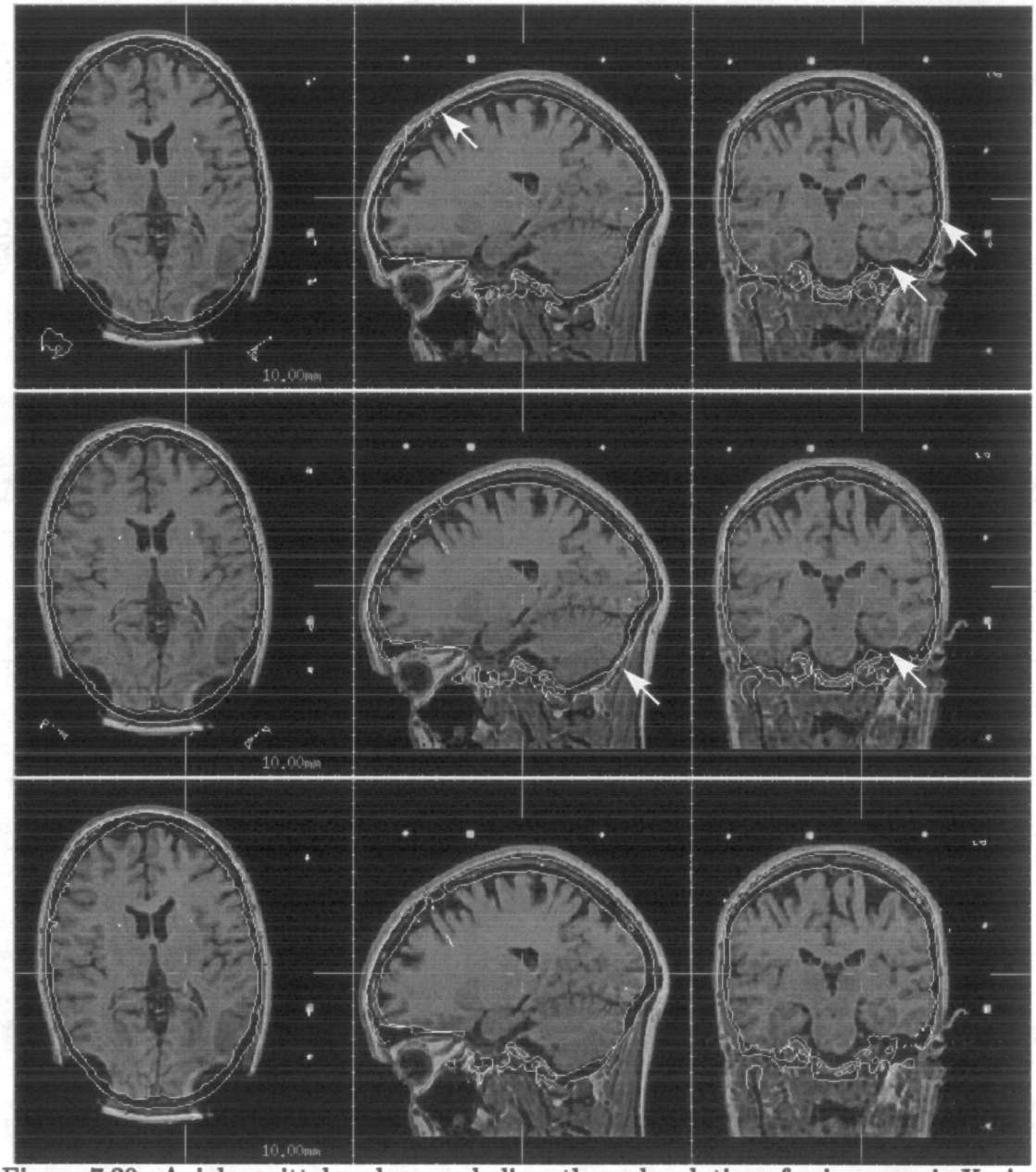


Figure 7.20: Axial, sagittal and coronal slices through solutions for image pair X: An example where the automated alignment including scale and skew (bottom) is superior to that using only a rigid transformation (middle) and the manual rigid estimate (top).

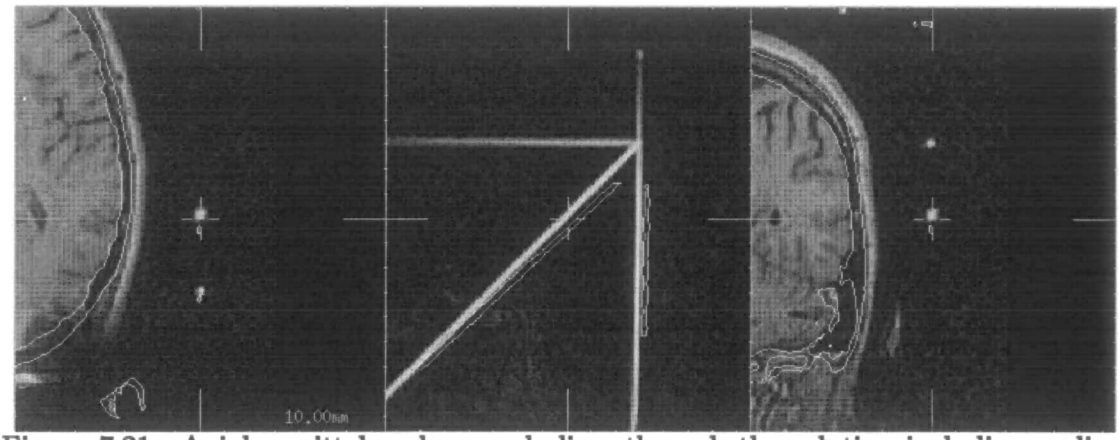


Figure 7.21: Axial, sagittal and coronal slices through the solution including scaling and skew for image pair X: Significant mis-alignment of stereotactic frame components indicating that the geometric distortion may be significantly worse further from the patient.

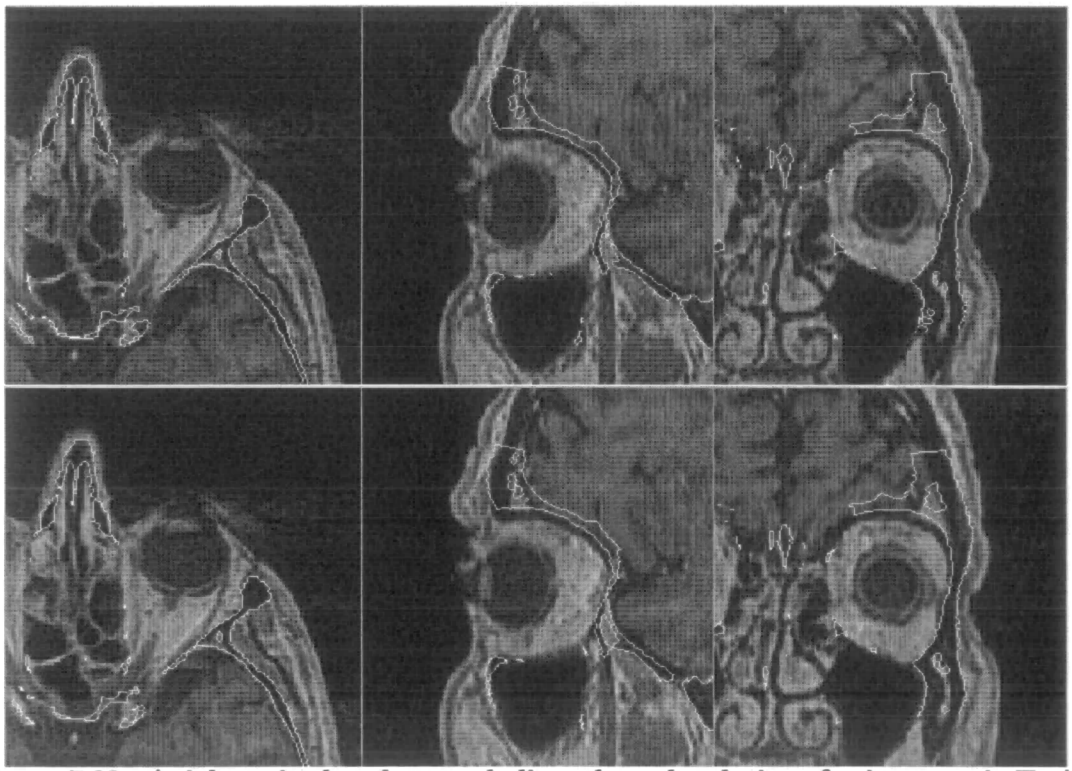


Figure 7.22: Axial, sagittal and coronal slices through solutions for image pair T: An example where the alignment including scale and skew parameters (bottom) and the rigid transformation estimate (top) are both visually mis-aligned.

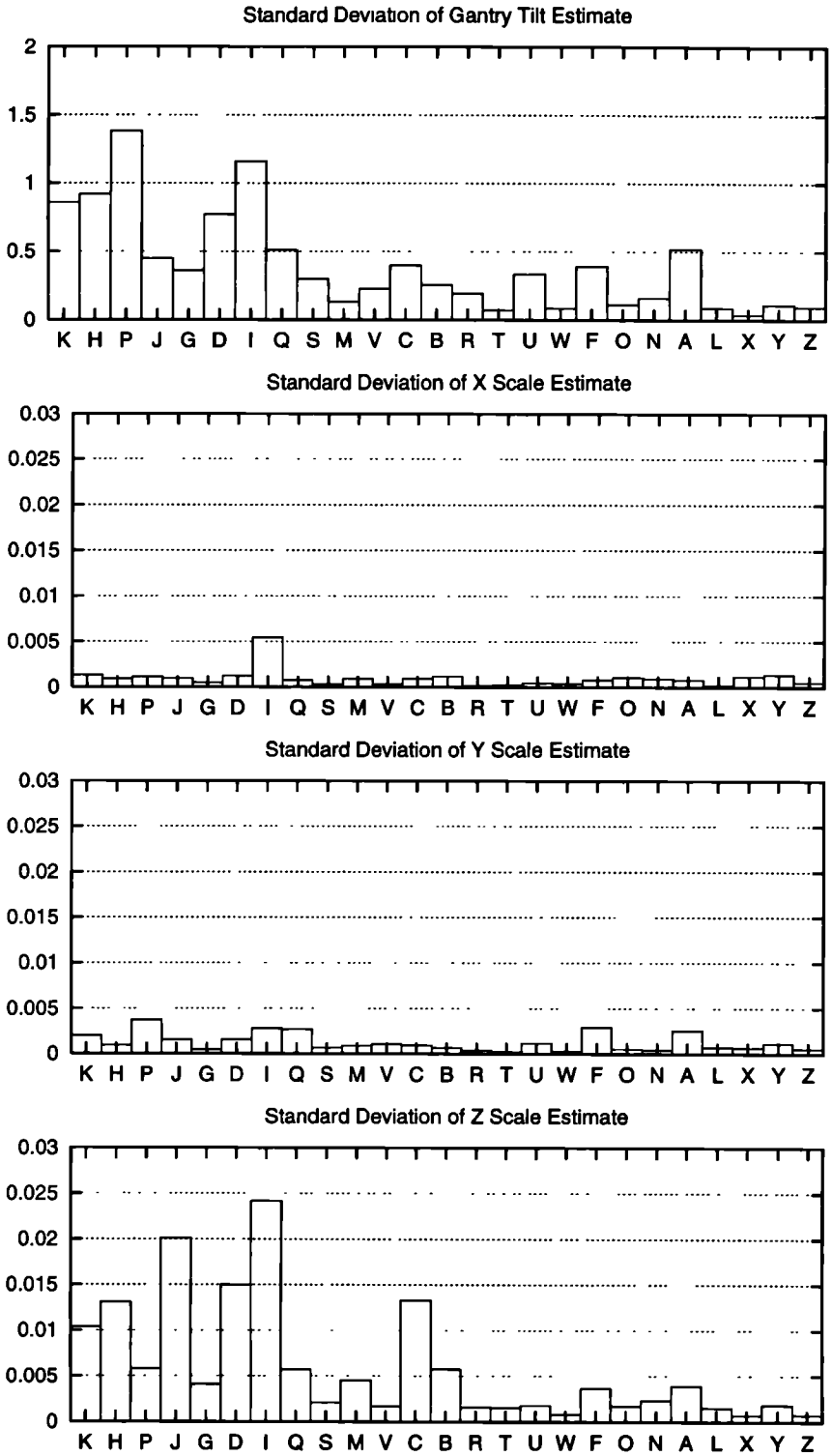


Figure 7.23: Standard deviation of orthogonal scaling and gantry tilt estimates for each image pair (from 16 starts), ordered in terms of increasing (left-to-right) minimum axial extent of the data.

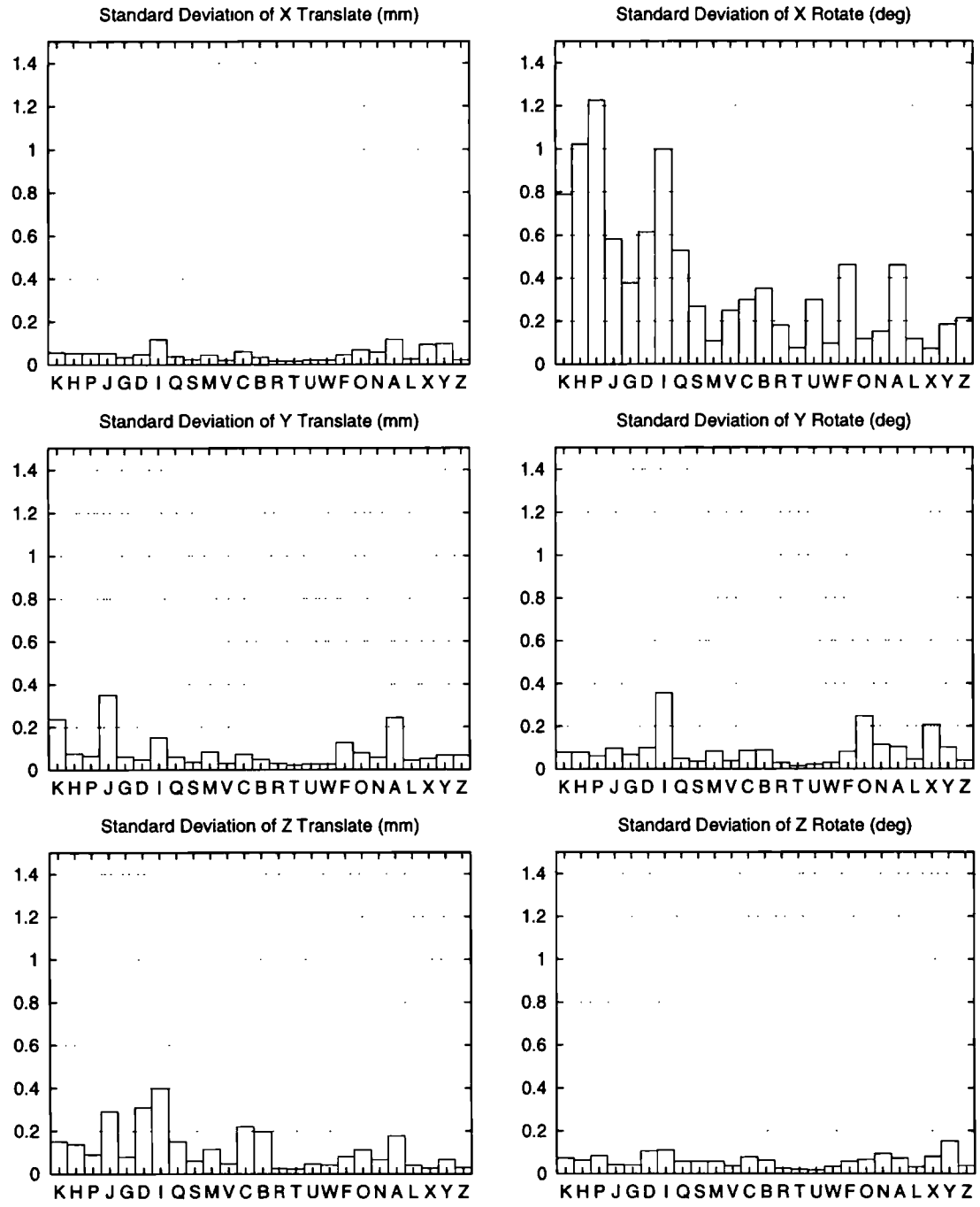


Figure 7.24: Precision of estimates of the translation and rotation parameters for each clinical image pair when optimising rigid, scale and gantry angle parameters, ordered in terms of increasing (left-to-right) minimum axial extent of the data.

7.7 Summary

In this chapter we have attempted to look at the application of global voxel similarity to truly retrospective MR-CT image registration. At the beginning of the chapter the relationship between noise, spatial resolution and registration accuracy was investigated on an 'ideal' dataset. We then went on to look at the practical limits of alignment accuracy on a clinically representative image database. We have seen that we can fully automatically recover a clinically useful rigid registration estimate from a wide range of image data, with an accuracy visibly comparable or superior to that provided by current manual point based techniques. The accuracy of the rigid registration in many practical cases is dominated by the geometric integrity of the image data.

By extending the registration to 10 parameters we can usefully recover or validate estimates of the image scaling and CT gantry tilt parameters. In many cases this can provide an appreciable improvement in registration accuracy and is a valuable check of the validity of the image geometries. Multi-start experiments indicate though that there can be an appreciable decrease in the precision of the transformation estimate when extra parameters are included. This suggests that such an approach may only be useful where large volume acquisitions are to be registered. Here the effect of scaling and skew errors is greatest on the overall registration, while the precision of the 10 parameter estimate is greatest.

An extension of this work has looked at the application of a nine-parameter (rigid registration with orthogonal scaling) registration to prospectively calibrate MR images using a phantom. This provides actual scaling estimates as opposed to relative scaling between the MR and CT images which can be important for surgery guidance applications. The limitation of this calibration approach is that it is only applicable to images acquired with using the calibration imaging protocol, and cannot be applied to data imported from other hospitals. In such cases rescanning locally would result in extra cost and in cases of CT, increased X-ray dose to the patient.

On this selection of data there were a significant number of examples where misalignment was visually apparent even with the inclusion of scaling and skew in the registration. These cases were not due to a failure in the registration algorithm to correctly recover these parameters, but were most attributable to more complex geometric distortion in MR. The other main factor limiting the accuracy of the final registration is the presence of motion artefact in both the MR and CT. This is a big unknown and its form is very dependent on the type of motion and the mode of acquisition used in imaging.

Chapter 8

Retrospective Alignment of Clinical MR and PET of the Brain

8.1 Introduction

Manual point based registration has been employed over the last few years to align MR and PET brain studies at our site. This chapter looks at the application of voxel similarity to automate this registration process. In chapter 5 normalised mutual information was identified as providing a robust recovery of registration on a database of MR-PET studies, which included an independent estimate of registration. This chapter looks at the application of normalised mutual information to align a range of typical clinical MR and PET data acquired at our site. The aim is to completely automate the registration process and, in addition provide improved accuracy over the current approach.

As with the previous chapter we start by looking experimentally at the fundamental limits on accuracy imposed by noise and spatial resolution of the image data. There is then a brief overview of the range of image characteristics found within the database of images used for the validation experiments. Although an independent *gold standard* estimate is not available for this data, we can compare the automated estimates to conventional manual point based estimates. In addition, visual inspection of the registration estimates is employed to assess the final accuracy. As with the MR-CT experiments in the previous chapter, the precision of the estimates is assessed using a simple multi-start experiment for each of the image pairs tested.

8.2 The Effect of Image Noise and Resolution on Alignment Accuracy

This set of experiments is aimed to investigate how the image noise and sampling resolution limits the accuracy of the automated estimate for MR-PET registration. To do this we are going to take one of the images from the Vanderbilt study (patient 5) with an accurate independent estimate of alignment, and change the noise level and spatial resolution.

8.2.1 Experimental Method

Image Sampling Resolution

Essentially we can take the same approach as that described in section 7.2.1. The original spatial resolution of the PET image is very much lower than that of the MR or CT. As a result we are going to look at a range of lower cubic resolutions of between $8.0mm$ (equivalent to the PET slice thickness) and $1.0mm$ with steps of $0.5mm$. Using the same sampling scheme as described in chapter 4, linear interpolation is used to increase resolution and Gaussian blurring is used to reduce resolution from the original image sampling rate. Example slices through images at two resolutions are shown in Figure 8.1.

Image Noise

As with the MR-CT estimates in the previous chapter, a range of Gaussian noise signals were added to both the MR and PET values at the original image resolution. The level of the Gaussian noise added to both images was set to varying fractions of the range of values in the images from 0% to 100% in steps of 10%. Examples of the resulting images are shown for two levels in Figure 8.2. As with the MR-CT experiments a simple estimate of the noise level in the original image data was carried out. The range of values in the original MR was 1746 and the values recorded in a region of air in the image had a standard deviation of around 10 units (0.57% of the range of values). The range of values in the original PET image was 795 and the values measured in a region of air in the image had a standard deviation of 23 units (2.9% of the range of values).

As with image resolution, in order to look at both the accuracy of a final estimate and its precision with respect to the marker estimate, 16 random starts were used at each noise level, and the MIPVD (section 6.4.1) evaluated for each estimate.

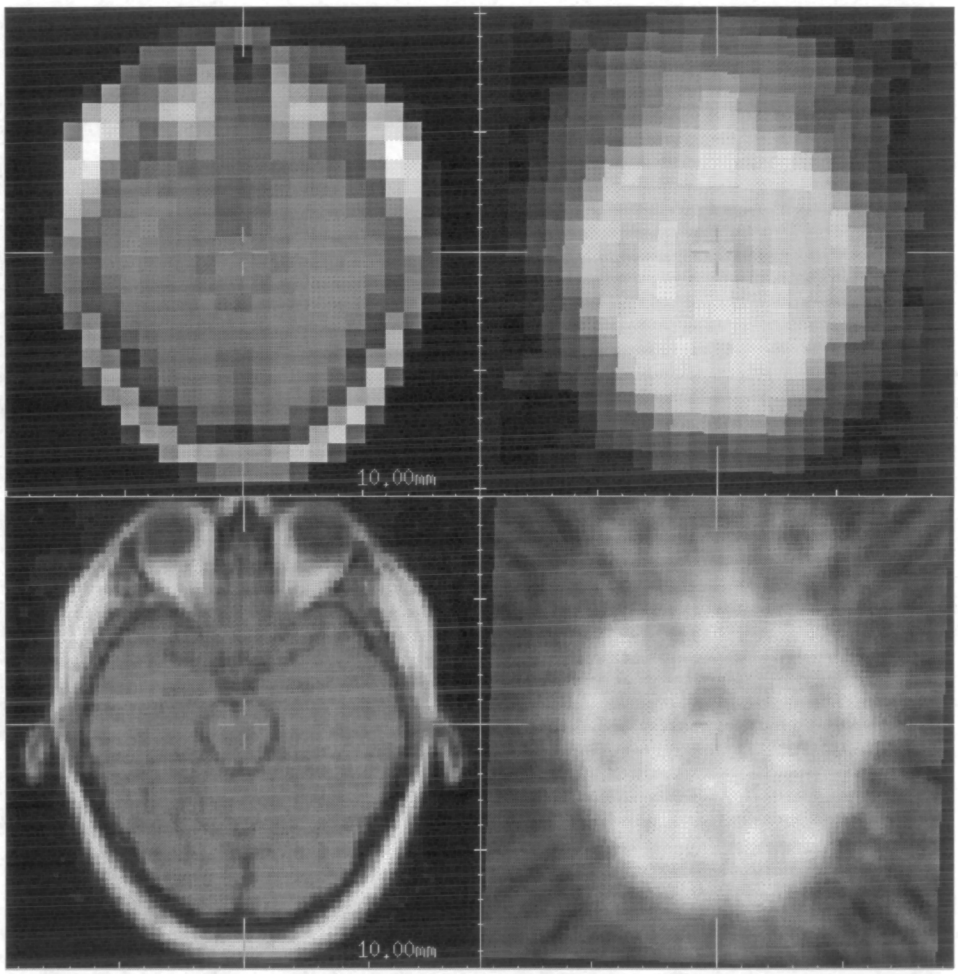


Figure 8.1: Examples of two image resolutions used to investigate registration accuracy, left MR , right PET. Upper image is with $8mm$ cubic voxels and lower image is with $2mm$ cubic voxels.

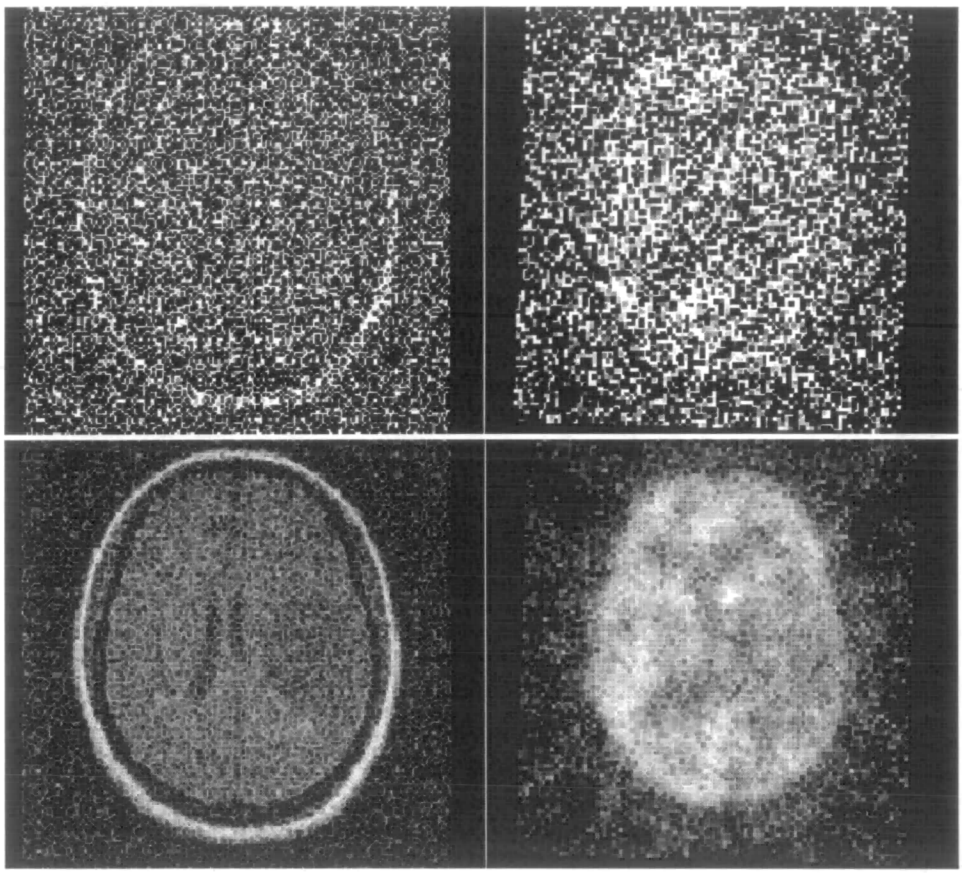


Figure 8.2: Examples of two levels of additive Gaussian noise used in experiments, left MR, right PET. Upper image is with added Gaussian noise with standard deviation equal to 80% of the data range. Lower image is with added noise at 10% of the data range.

8.2.2 Results: Image Resolution

The errors with respect to the marker estimate, from 16 random starts at each image resolution are summarised by the graph of the averaged error for these starts in Figure 8.3. This shows an overall fall in voxel misalignment as the sampling resolution of the images is increased. At the highest image resolution ($1.0mm$) the mean voxel displacement averaged over 16 starts between the automated and marker estimates is $2.08mm$ with a standard deviation of $0.26mm$ over the imaged patient volume. This is significantly worse than the MR-CT estimates due to the lower spatial resolution of the PET imagery.

As a comparison, the experiment was repeated using a pair of locally acquired clinical image data. Here the PET slice thickness is only $3.75mm$ (compared to the $8mm$ of the Vanderbilt data) and the spatial resolution is approximately $8mm$ FWHM (compared to $15mm$). The MR spatial resolution was also higher with voxel dimensions of $0.859 \times 0.859 \times 1.2mm$. The resulting graph of error with respect to a manual point based estimate is shown in Figure 8.4. This shows a surprisingly small MIPVD between the manual point based estimate and the automated estimate at the highest resolution ($0.57mm$) with a standard deviation of only $0.053mm$. There is a distinct peak at the $4mm$ resolution which may be as result of the simple tri-linear interpolation used.

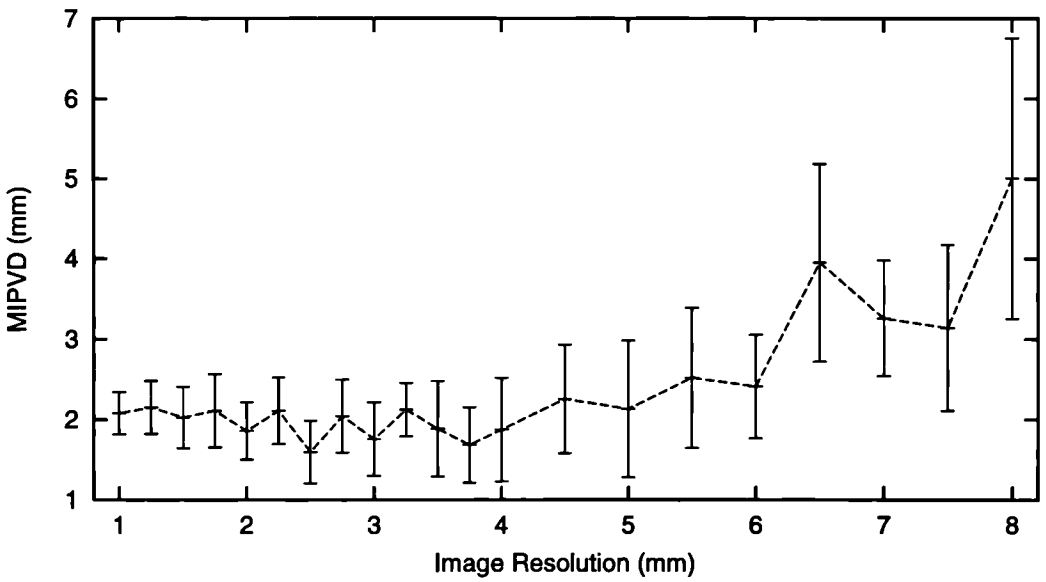


Figure 8.3: Vanderbilt Image Pair 5: MIPVD between the estimate provided by normalised mutual information and the marker estimate as sampling resolution is varied: Average with standard deviation (error bars) over 16 starts at each sampling resolution.

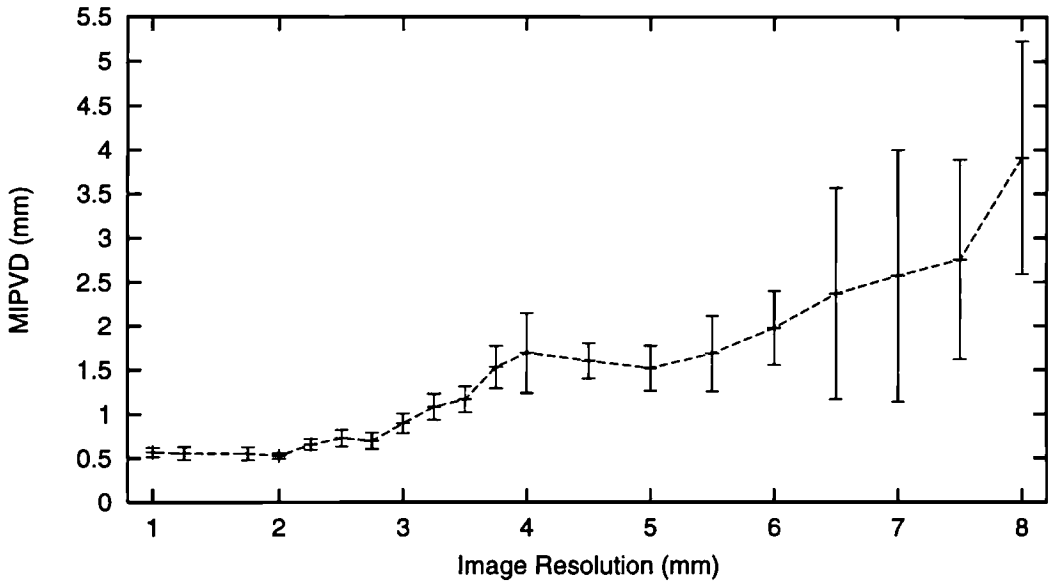


Figure 8.4: Local Image Data: MIPVD between the estimate provided by normalised mutual information and a manual point based estimate as sampling resolution is varied: Average with standard deviation (error bars) over 16 starts at each sampling resolution.

8.2.3 Results: Measurement Noise

The average MIPVD (section 6.4.1) of estimates at each level of noise, along with their standard deviations are shown in Figure 8.5. As the noise level is increased there is an overall fall in both the accuracy and the precision of the registration estimates with respect to the marker estimate. At levels up to around 20% of the range of the image values, the effect on accuracy is relatively small. This is probably due to the low resolution of the PET data and also the higher levels of noise already present in the PET image. There is some indication of peaks in both the MIPVD and its deviation at levels of 40% and 80% of the range of image values, but these are small compared to the standard deviation of the MIPVD at these levels.

8.2.4 Discussion

The experiments in this section have showed that for MR-PET brain image registration the alignment accuracy is considerably lower than that for MR-CT alignment. In particular the limited spatial resolution of the PET image appreciably limits the final accuracy. Evaluation of the image alignment at sampling resolutions higher than around $2mm$ may not significantly improve the final registration accuracy on data acquired locally.

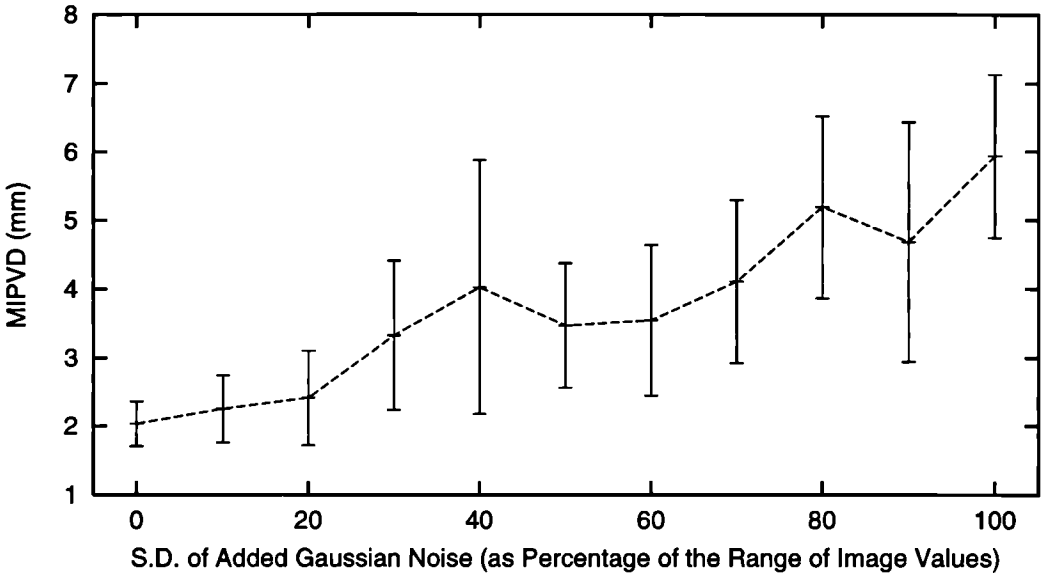


Figure 8.5: Vanderbilt Image Data: MIPVD between estimate provided by normalised mutual information and the marker estimate as noise level is varied: Average with Standard Deviation (error bars) over 16 starts at each noise level.

8.3 A Database of Clinical MR-PET Images

8.3.1 An Overview of Clinical Image Characteristics

The basis for this set of registration experiments are a set of 14 clinically acquired image pairs of the brain. The first 10 image pairs (A-J) are representative of the majority of MR-PET cases at our site. There are no space occupying lesions in the data and the images are generally acquired for research studies into functional abnormalities in cases such as epilepsy. An example is shown in Figure 8.6. This and all the other example images in this section are shown with manual registration estimates (described in the following section).

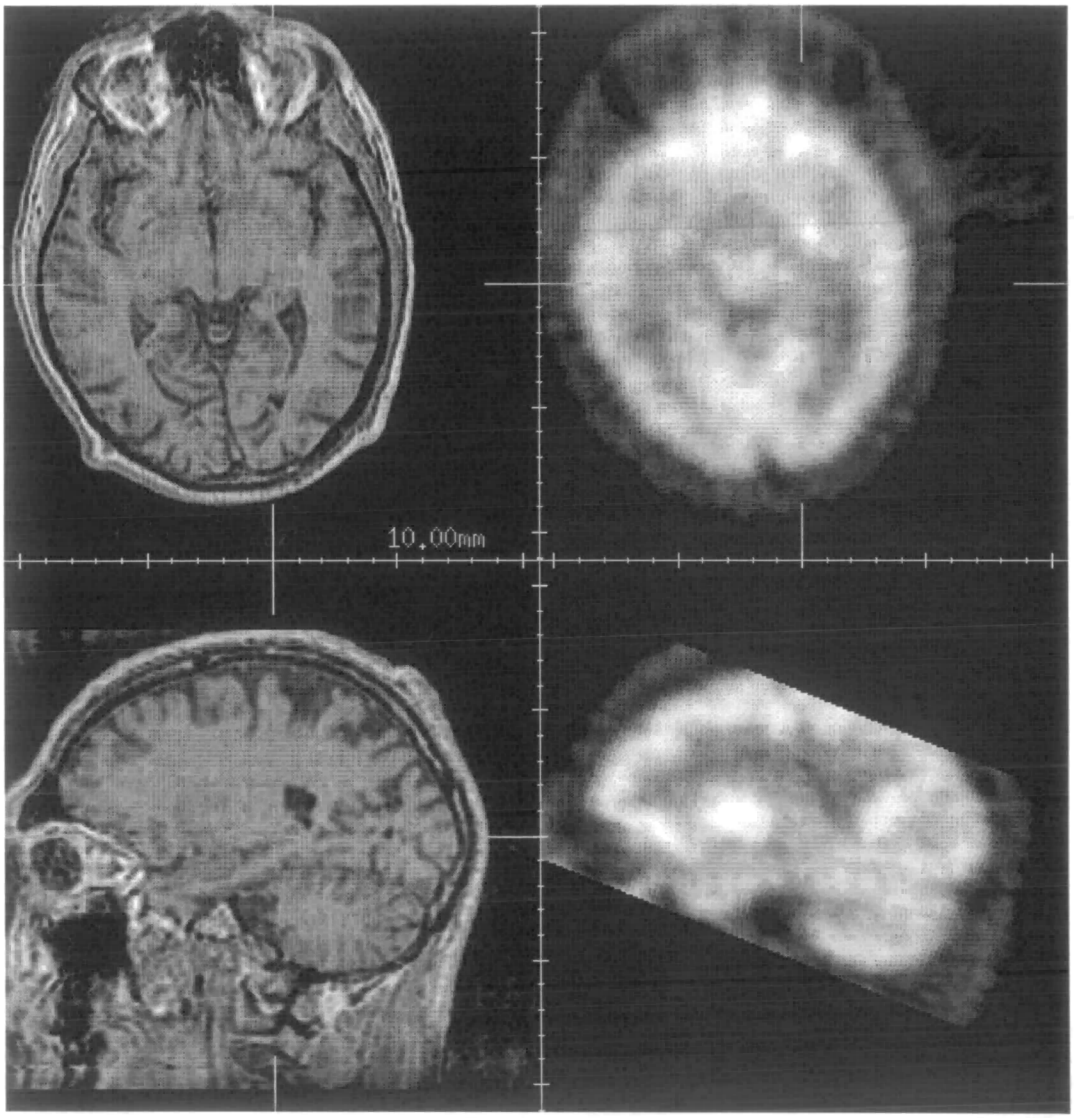


Figure 8.6: Example axial (top) and sagittal (bottom) slices through MR (left) and PET (right) images of patient C showing normal anatomy.

In addition to these, the last four image pairs (K-M) are included as more extreme cases of anatomy and physiology. The images for patient K (figure 8.7) display significant atrophy of the brain resulting in a considerably more complex brain surface structure. This is reflected in the higher resolution MR image, but because of its limited spatial resolution, is not as evident in the PET image. Patient M (figure 8.8) illustrates signifi-

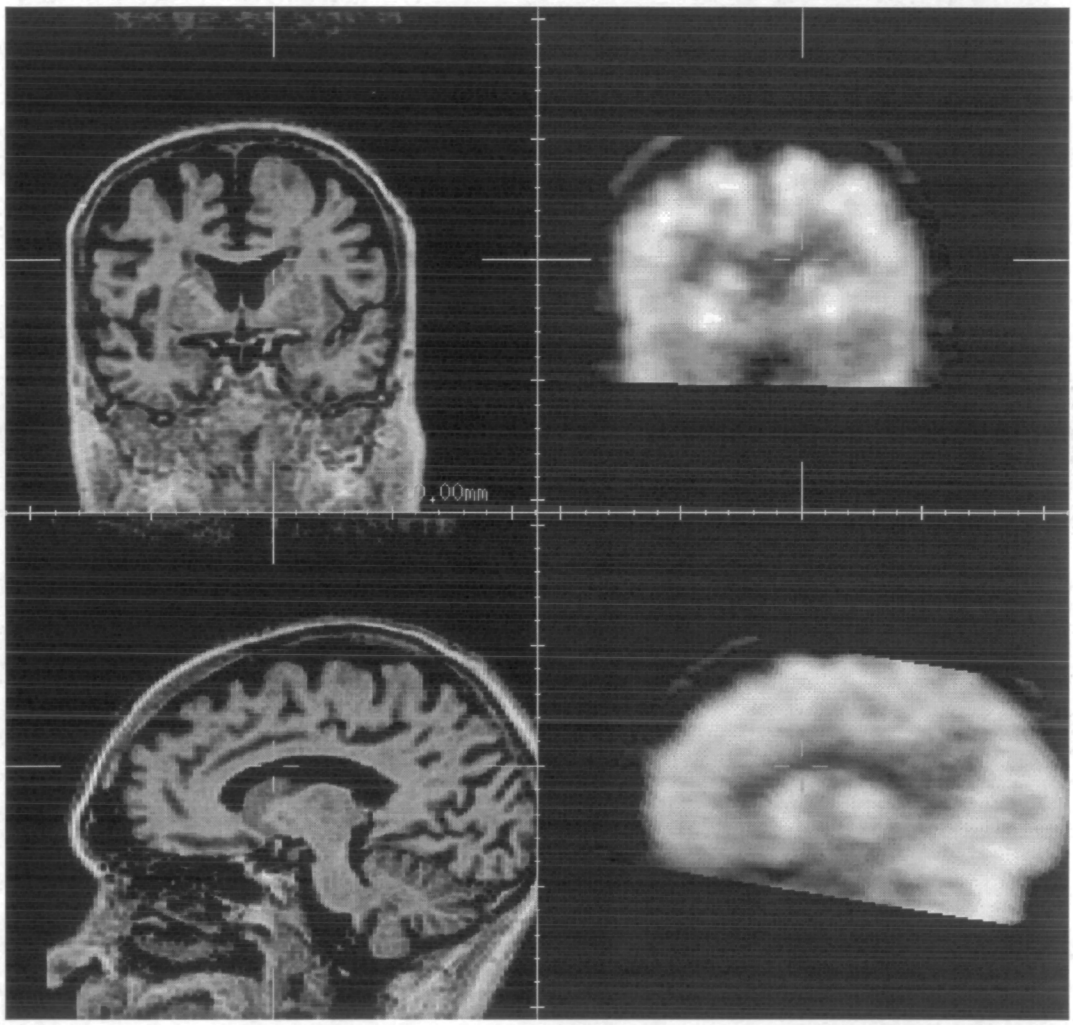


Figure 8.7: Example coronal (top) and sagittal (bottom) slices through MR (left) and PET (right) images of patient K showing significant atrophy.

cantly abnormal anatomy of a child resulting from the growth of a tumour. This poses a particular problem because of the significantly reduced uptake in the PET image which will alter the balance of the values in the joint probability distribution. Patient M and also L illustrate the anatomy and uptake in the brain of a child, where the thickness of the scalp and skull is significantly less than the more commonly imaged adult brain. This may also influence the registration measure and the complexity of the resulting

parameter space.

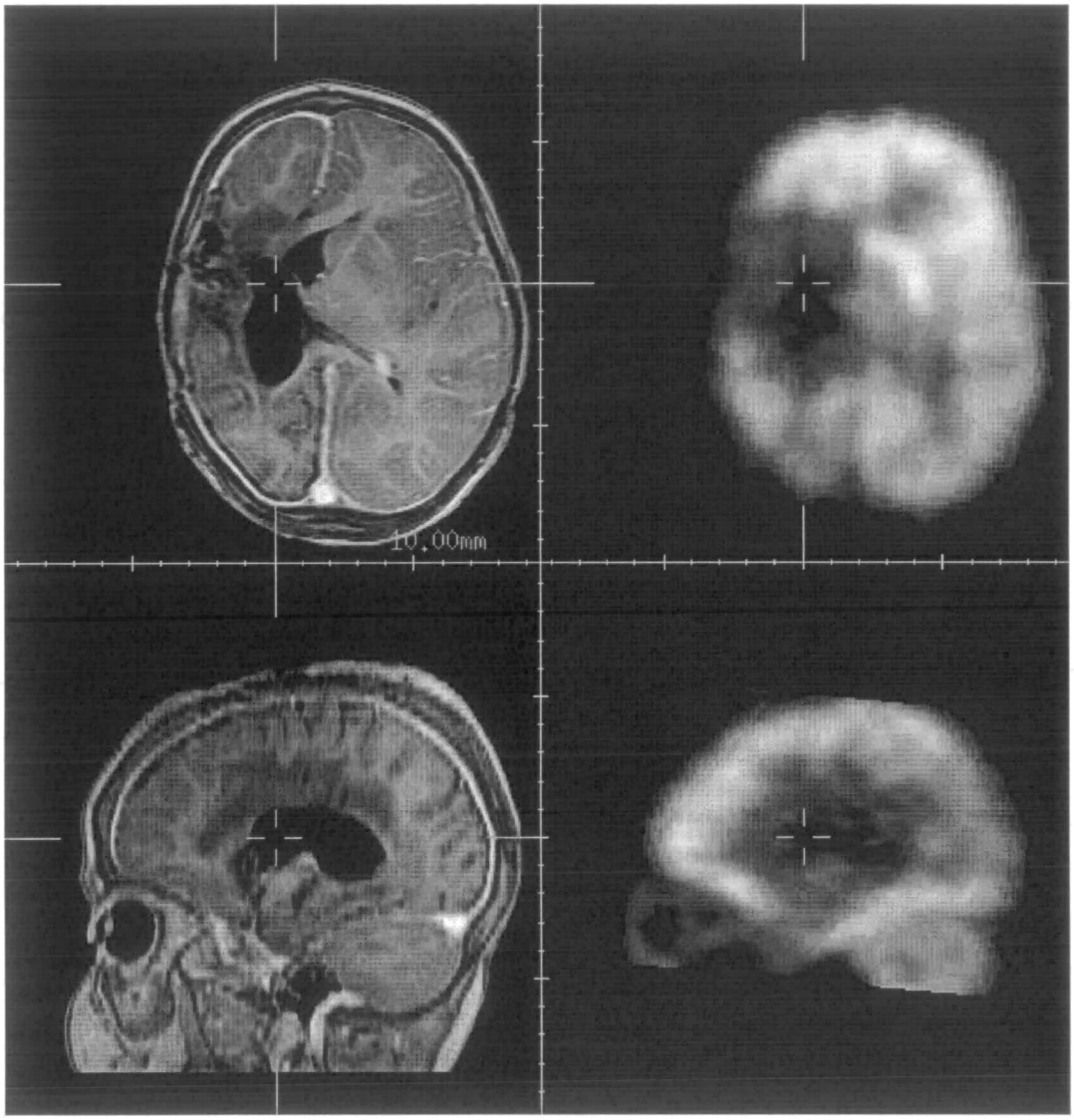


Figure 8.8: Example axial (top) and sagittal (bottom) slices through MR (left) and PET (right) images of patient M showing significantly abnormal anatomy and tracer uptake in the brain of a child.

Image set N (figure 8.9) is an example of an alternative PET tracer image acquired at the M.R.C. Cyclotron Unit at the Hammersmith Hospital in London. It shows the 3D PET reconstruction image of ^{11}C Temozolomide uptake, which is a tracer used to localise astrocytomas (visible on the patients right in the axial slice). There are considerably fewer anatomical details in this PET image and appreciably higher levels of noise in comparison to the earlier clinical FDG images.

All data consists of nominally transaxial slices but with a range of orientations typical in routine clinical scanning. All the PET images were acquired on a Siemens/CTI

scanner (Knoxville, Tennessee, USA). The PET images were all static, summed from axially smoothed dynamic ^{18}FDG acquisitions reconstructed to 31 slices of 128 by 128 voxels, except for image pair N which was a 3D reconstruction. The voxel size was $2.0 \times 2.0 \times 3.375mm$ and the point spread function has a full width half maximum of approximately $8.0mm$. This is appreciably higher resolution than the PET data used in the Vanderbilt registration study. The MR acquisitions came from 3 different scanners, one 2D spin echo image (patient A) from a 1.5T Philips Gyroscan S15/HP, 5 3D gradient echo images (patients B to F and K to N) from a 1.5T GE Signa, and 4 3D gradient echo images (patients G to J) from a 1.5T Philips Gyroscan ACS II. All were T1 weighted and intended to show good grey/white matter delineation. Large MR volumes were used to give a good range of tissue types (white matter, grey matter, skull, scalp etc.) for registration. Table 8.1 shows the field of view and voxel dimensions of the individual MR datasets.

MR Image Volume Parameters

Image Pair	Voxel Size (mm)	F.O.V. (mm)
A	0.86 × 0.86 × 2.5	220 × 220 × 137.5
B	0.90 × 0.90 × 1.5	230 × 230 × 186
C	0.94 × 0.94 × 1.5	240 × 240 × 186
D	0.94 × 0.94 × 1.5	240 × 240 × 186
E	0.94 × 0.94 × 1.5	240 × 240 × 186
F	0.94 × 0.94 × 1.5	240 × 240 × 186
G	0.90 × 0.90 × 1.2	230 × 230 × 144
H	0.86 × 0.86 × 1.2	220 × 220 × 144
I	0.86 × 0.86 × 1.2	220 × 220 × 144
J	0.86 × 0.86 × 1.2	220 × 220 × 144
K	0.98 × 1.72 × 0.98	250 × 250 × 250 *
L	0.86 × 0.86 × 1.5	220 × 220 × 150
M	0.82 × 0.82 × 2.5	210 × 210 × 167.5
N	1.0 × 1.0 × 1.0	142 × 250 × 254*

Table 8.1: MR image volume parameters. All images nominally axial except those marked with * which were reformatted to be axial.

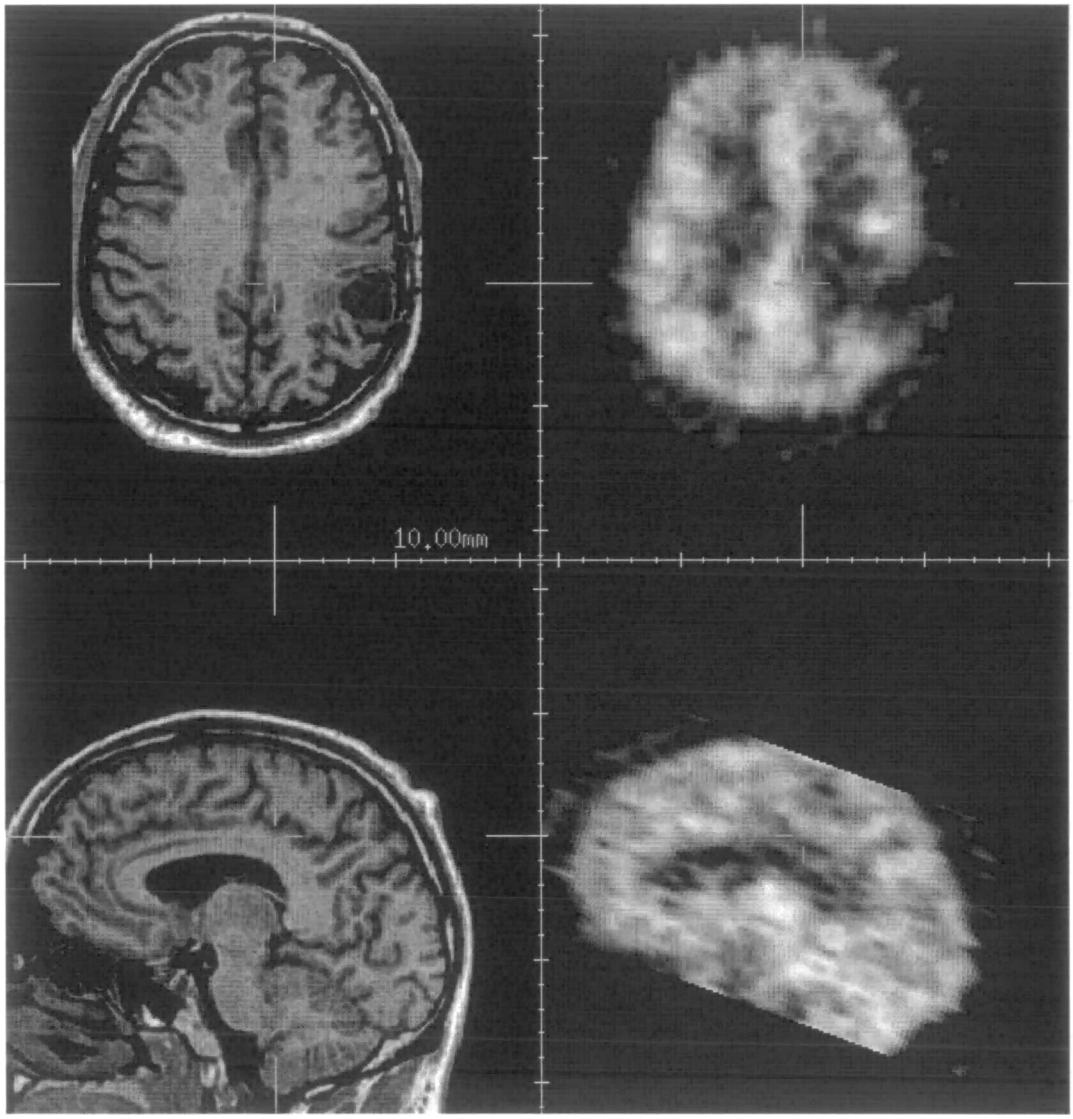


Figure 8.9: Example axial (top) and sagittal (bottom) slices through MR (left) and PET (right) images of patient N showing ^{11}C Temozolomide PET tracer uptake.

8.3.2 An Independent Manual Estimate of Registration

Image pairs A-M were manually registered by interactive location of between 10 and 14 corresponding anatomical point landmarks [32]. To check registration accuracy the registration estimates were inspected interactively using the *rview* software to confirm alignment [104] was comparable to that used in clinical practice. Because of the use of an alternative PET tracer in image pair N which delineates significantly less anatomy, it was not possible to identify corresponding landmarks to manually register this image with the MR image. In this case a manual interactive registration using *rview* (section 6.4.2) was carried out by manipulating the 6 rigid parameters until a visually acceptable solution was provided. All the estimates are shown in table 8.2 in the form of 3 translations and 3 rotations to map each PET image to MR coordinates (where the x-axis is the patient right-left, the y-axis from front to back and the z-axis feet to head). All registrations were confirmed visually to be within about $3mm$ over the brain volume.

Manual MR-PET(^{18}FDG) Registration Estimates

Patient	Translation (mm)			Rotation (deg.)			R.M.S.	Num.
	t_x	t_y	t_z	θ_x	θ_y	θ_z	Error	Points
A	2.9	23.3	8.4	11.7	-1.9	-0.5	3.9	12
B	-2.6	1.8	7.6	24.2	1.1	8.7	3.6	12
C	-9.6	-9.8	22.8	21.6	-0.8	6.1	3.0	14
D	-3.5	7.0	15.4	16.1	-2.7	3.8	2.9	14
E	-1.8	18.3	0.6	11.3	0.2	0.6	2.7	12
F	-3.4	-11.0	0.0	18.4	-0.8	5.9	3.5	12
G	1.0	7.0	-8.7	18.6	4.3	-5.7	3.1	11
H	2.1	11.8	0.4	14.1	-0.9	-0.4	4.3	14
I	3.5	12.8	-7.3	18.3	0.5	3.7	2.9	10
J	0.7	22.6	-4.9	12.7	-3.0	2.3	2.5	11
K	-1.4	22.1	14.8	9.0	-0.8	0.5	4.2	10
L	0.1	-11.2	11.6	2.2	-6.7	-3.6	3.7	10
M	5.1	12.5	-0.2	5.3	1.9	2.7	2.7	10
N	-7.0	16.0	44.0	20.0	0.0	0.0	-	-
Max	5.1	23.3	44.0	24.2	4.3	8.7	4.3	
Min	-9.6	-11.2	-8.7	2.2	-6.7	-5.7	2.5	

Table 8.2: Transformation parameters estimated by manual point landmark identification (A-M) and interactive registration (N).

8.3.3 Common Types of Clinical Image Misalignment

The manual results provide a valuable indication of the scale of typical misregistration encountered between brain images in clinical practice. The largest translational misalignment is for patient N, in the between slice direction ($t_z = 44.0mm$). The MR image

slices in this case were initially sagittal and therefore the reformatted volume has a larger axial extent. In general, most of the MR acquisitions for brain studies with PET cover a significant portion of the brain and varying portions of the skull base and neck. In contrast the PET scan extent is limited and fixed to 31 planes and the imaged region is chosen to include as much of the brain as possible. The resulting translational differences between the two volumes can be large. As with CT, the main rotational difference is in the $\theta_x = 24.2°$ direction which corresponds to the nodding angle. This is due to the different design of head rest and different imaging protocols in the two modalities. As with MR-CT, the other two rotations are generally much smaller ($-6.7°$ and $8.7°$) and not biased in one direction.

8.4 Automated Rigid Registration

8.4.1 Algorithm Parameters

In order to ensure a fair test of automated alignment, identical parameter settings for the registration algorithm were used to align all the image pairs. The optimisation was carried out as described in section 5.6.

The choice of the number of bins to use for the evaluation of the measure was investigated with a simple experiment to examine the relationship between accuracy and bin size. The graph of MIPVD in Figure 8.10 shows the errors with respect to the marker estimate for the Vanderbilt image data used in section 8.2.1 as the number of bins is varied (using $1.5mm^3$ minimum voxel size) . This indicates that the choice of the number of bins used has little effect on the final registration accuracy. The number of bins used for these experiments was 64.

8.4.2 Experimental Design

As with previous experiments, a starting estimate of the centre of the imaged volumes aligned, with no residual rotations was used as the default starting estimate. Optimisation was applied from this estimate for all the image pairs and the results recorded to give an indication of robustness over the full image database. The translational and rotational components of the estimate for each image pair was then perturbed by $5mm$ and 5^o respectively to form 16 starting estimates, as described in the MR-CT experiments in section 6.6.1. Re-registration was then carried out form these starting estimates to provide an indication of the precision of the registration estimate.

In order to compare the manual and automated registration estimates the MIPVD between the two estimates was evaluated as described in section 6.4.1. As with the earlier

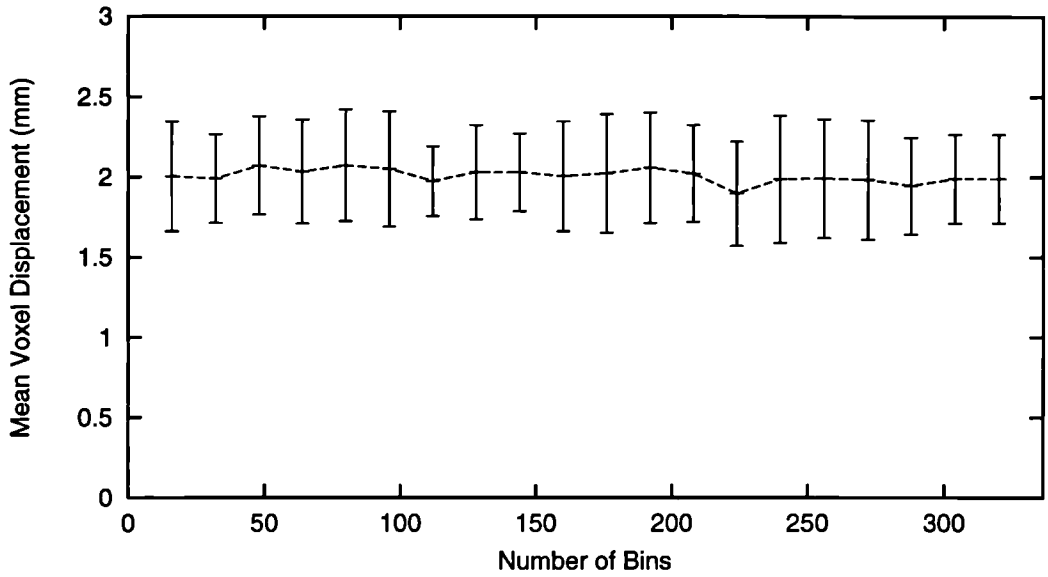


Figure 8.10: Vanderbilt Image Data: MIPVD between the estimate provided by normalised mutual information and the marker estimate as the number of intensity bins (for both images) in the joint probability distribution is varied: Average with standard deviation (error bars) over 16 starts at each number of bins.

MR-PET experiments a threshold of PET image intensity was interactively selected to define brain voxels in the PET images, in order to visually evaluate registration accuracy.

8.4.3 Results

Accuracy

The MIPVD (see section 6.4.1) between the manual and automated estimate for each image pair is illustrated by the bar chart in Figure 8.11. This shows the estimate with its standard deviation based on the 16 random starts. In all cases the precision indicated by the standard deviation of MIPVD is very much greater than the accuracy with respect to the manual estimate.

The automated method supplies estimates which are close to the manual registration estimate for all the datasets. All the estimates are within an MIPVD of $4mm$ of the manual estimate. There are four image pairs for which the mean patient voxel displacement is significantly greater than $3mm$ from the manual estimate. Inspection of these results indicate that the automated estimates provide visually superior registration compared to the manual estimates. The improvement provided by the automated estimates for image pairs K and N are illustrated in Figures 8.12 and 8.13. Points of misalignment between a brain surface threshold contour and the MR are indicated by arrows on the images. These indicate some y axis rotational error in the manual estimate for patient M, and x axis rotation (nodding) error in the estimate for patient K.

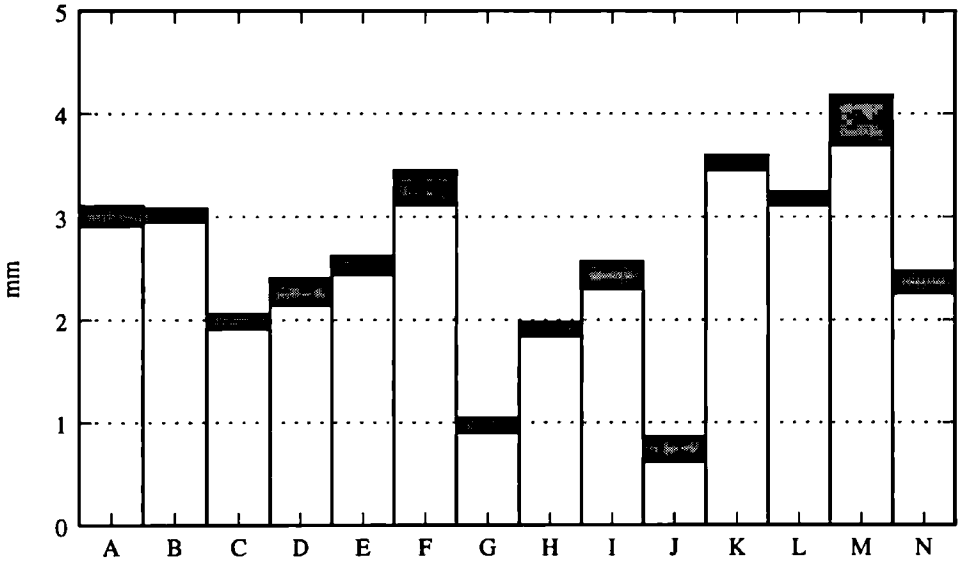


Figure 8.11: MR-PET: MIPVD in mm between rigid transformation estimates provided by manual alignment and automated optimisation of normalised mutual information (Standard deviation from 16 starts shown as error bars).

Precision

The standard deviation derived from 16 starts of each of the 6 rigid parameters are shown graphically in Figure 8.14. Overall, the standard deviation of each of the estimates is comparable. Patient M provides a slightly larger spread of results, particularly for the x axis rotation (nodding), the y axis translation and the z axis rotation. This is probably a result of the significant space occupying lesion present in this case (see Figure 8.8) which means that the brain boundary is poorly defined on the left of the image (patients right) in both modalities.

There is some indication that, over the full set of images the X and Y rotations (out of plane) have a greater spread of estimates. This may be a direct result of the limited of spatial resolution in the z axis, particularly in the PET image data. Other influencing factors may include the shape of the brain, and also errors in scaling (particularly in the MR) leading to a these parameters being poorly constrained.

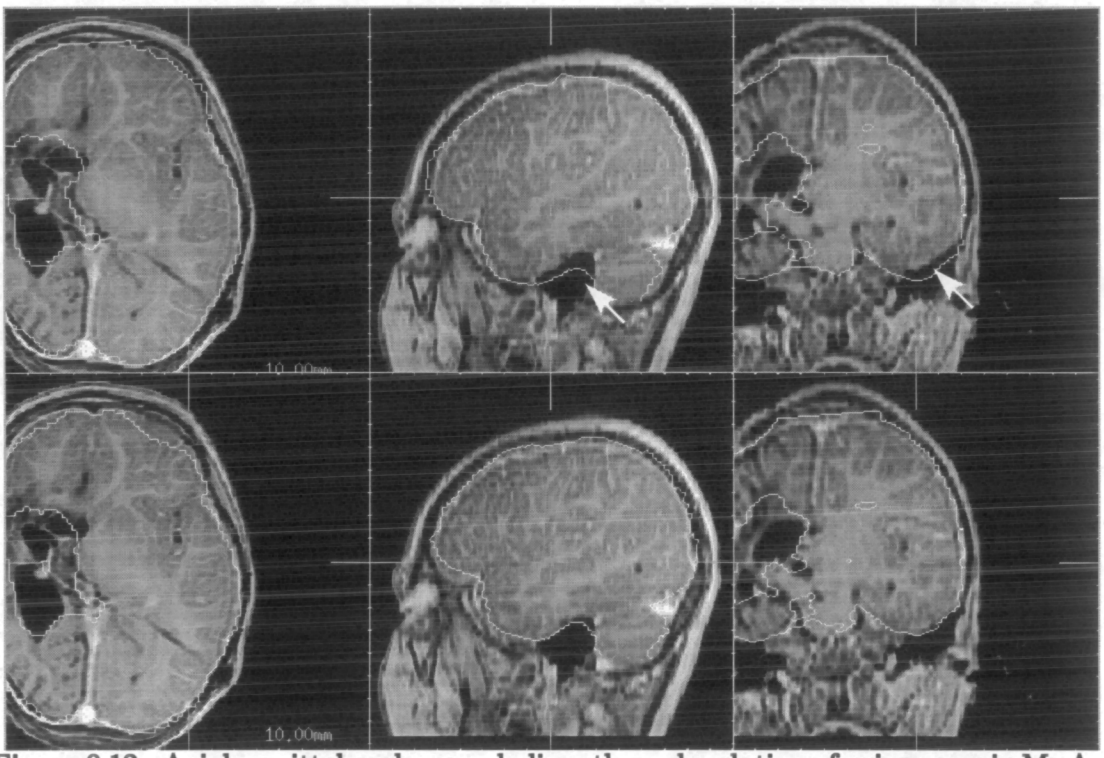


Figure 8.12: Axial, sagittal and coronal slices through solutions for image pair M: An example where the automated alignment (bottom) is visibly superior to the manual estimate (top).

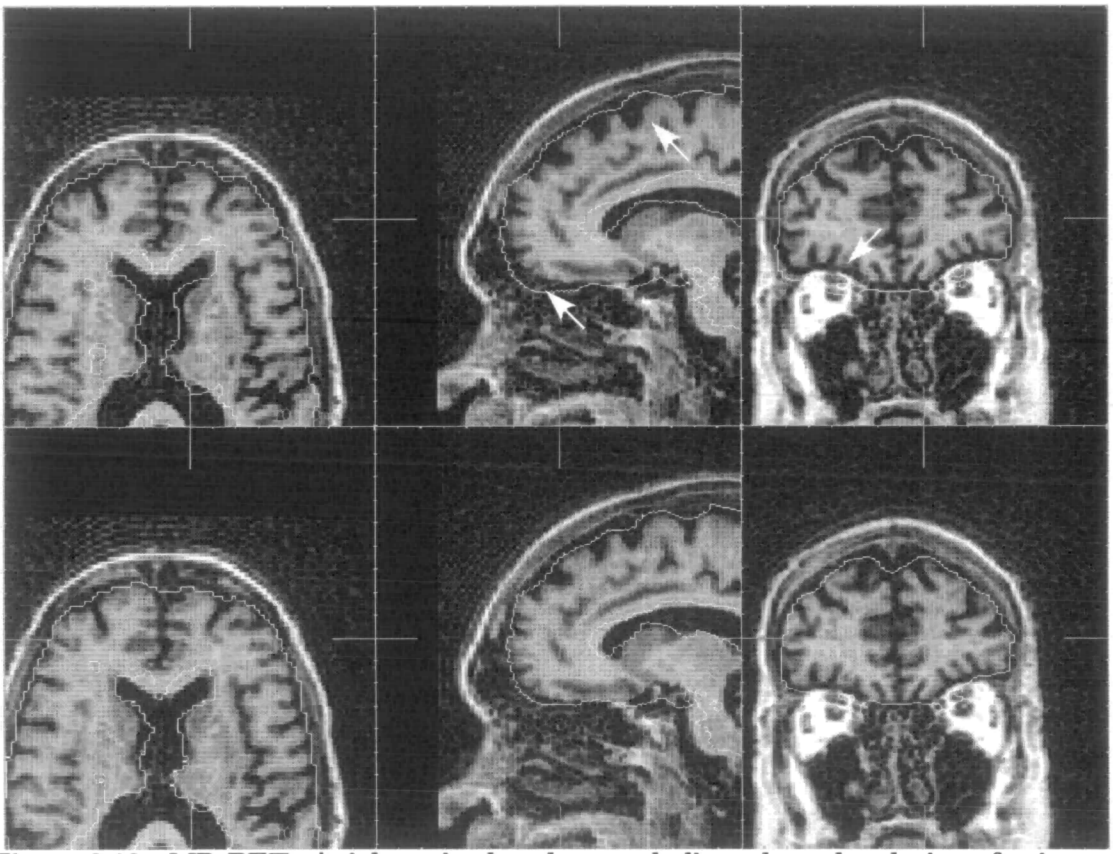


Figure 8.13: MR-PET: Axial, sagittal and coronal slices through solutions for image pair K: An example where the automated alignment (bottom) is visibly superior to the manual estimate (top).

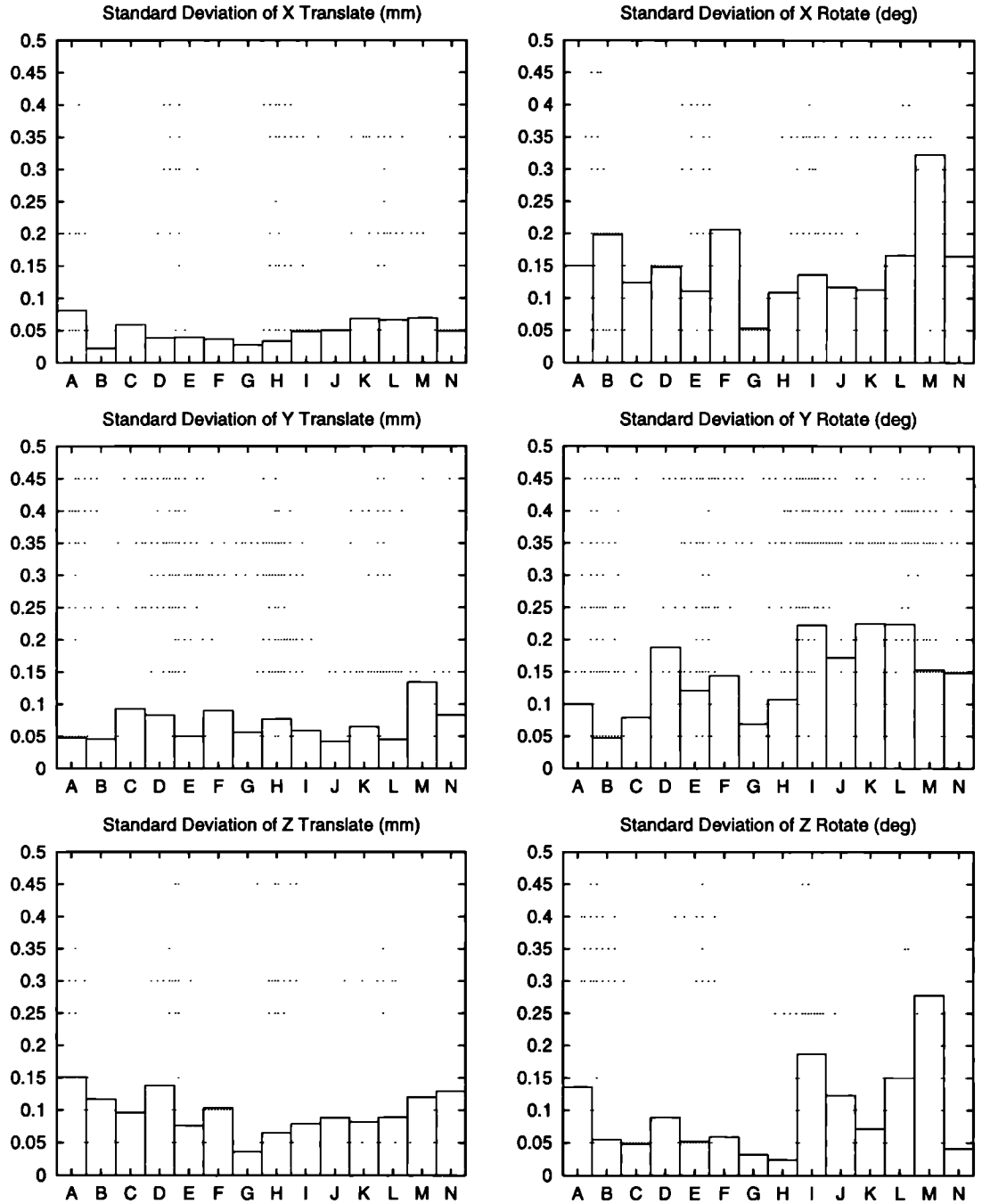


Figure 8.14: MR-PET: Precision of estimates of rigid registration parameters for each clinical image pair, when optimising rigid registration parameters.

8.5 Summary

This chapter has assessed the applicability of multi-resolution optimisation of normalised mutual information to retrospective alignment of MR and PET brain images. Over a database of typical data (including a range of image content and image misregistration) it is possible to fully automatically recover alignment. Numerical differences between the manual and automated estimates are within the expected accuracy of the manual point based technique. The automated estimates are visually comparable and often superior to those provided by manual point based registration.

The final registration accuracy for most cases is dominated by the limited spatial resolution of the PET imagery. Although MR scaling (as well as motion during the MR and PET scans) will still contribute to registration errors, their impact on the alignment may well be small in most cases, compared to the errors due to spatial resolution and noise. This is different to the MR-CT case where alignment accuracy may be dominated by the geometric integrity of the data.

In these experiments, unlike those for MR-CT, no attempt was made to recover scaling and skew estimates for the data. PET images are not acquired with a gantry skew angle, but scaling errors may still contribute, particularly for the MR images. Initially, experimentation with registration incorporating scaling indicated that it was not possible to robustly recover scaling estimates between MR and PET images. This almost certainly is a result of the limited spatial resolution of the PET images. Further work on developing more accurate independent estimates of registration is required before we can address the effects of MR distortion.

Chapter 9

The Alignment of Images With Severe Intensity Distortion

9.1 Introduction

For many types of medical image data we can assume that a material will exhibit roughly the same measurement at different locations within the image. This though, is sometimes not the case and for some classes of medical image data the sensitivity of the imaging device varies considerably over the field of view. These images still provide clinical information which is usefully related to other modalities.

In cases of severe measurement (intensity) distortion, the fundamental assumption that there are corresponding regions of uniform values within the images is broken. The co-occurrence of the most probable values in the two images does not correspond to true alignment. As a result, measures such as those investigated in the thesis so far do not provide a true indication of alignment.

This chapter introduces an extension of entropy based registration measures to allow the alignment of an image from a modality containing intensity distortion, to one with little distortion, without the need to separately correct the distorted image. This makes use of a three source measure of information to include spatial encoding of measurement values. The technique is applied to the alignment of distorted MR with PET images of the pelvis and surface coil MR with CT and head coil MR images of the brain.

9.2 Measurement Inhomogeneity

It is possible, particularly with some types of MR image, that the measured value of a physical property is affected by the location of that measurement within the field of view of the image. In general, the actual measurement $m(\mathbf{x})$ is some function $G()$ of both the value exhibited by a material $\tilde{m}(\mathbf{x})$ and its location $\mathbf{x}$,

$$m(\mathbf{x}) = G(\tilde{m}(\mathbf{x}), \mathbf{x}). \tag{9.1}$$

In MR imaging this is due to R.F. field B_1 inhomogeneity, and commonly results from the receiver coil [66] when imaging regions with a larger field of view such as in the abdomen and pelvis. The problem is particularly severe when imaging with a surface coil. A secondary cause of variation in intensity across an MR image is the effect of filtering of data prior to Fourier reconstruction resulting in 'windowing' of the image data in the spatial domain.

In MR these distortion affects can generally be modeled [13] as a variable attenuation factor $G_M(\mathbf{x})$ so that the measurement is given by

$$m(\mathbf{x}) = G_M(\mathbf{x})\tilde{m}(\mathbf{x}). \tag{9.2}$$

9.2.1 Measurement Similarity and Inhomogeneity

The most common task where intensity distortion poses a problem in multi-modality alignment, is where one image contains severe intensity distortion, while the second does not. An example is when comparing structures in an MR surface coil image $m(\mathbf{x})$ with a X-Ray CT image $n(\mathbf{x})$. In such a case, when we evaluate a measure such as mutual information, the mapping between intensities $\phi : M \mapsto N$ varies depending on the location $\mathbf{x}$.

There is a considerable body of work devoted to the problem of retrospectively correcting the intensity distortion in MR images [4, 13, 60, 101]. A solution to this problem is particularly important for intensity based segmentation techniques [9] which are heavily dependent on uniform sensitivity over the image. In effect, intensity distortion is an *amplitude modulation* of the image data. Correcting the distortion can be compared to amplitude demodulation of the image signal from a non-periodic carrier of unknown form. In practice these correction techniques must employ assumptions about the nature of the distortion function $G_M(\mathbf{x})$ to be able to separate intensity variation due to changes in sensitivity and intensity variation due to the structure of the objects being imaged.

One approach to solving the registration problem is simply to apply an intensity correction algorithm to the distorted MR image prior to aligning it with another modality.

This approach though makes no use of the additional information available from the second modality. Any alignment parameters derived from an independently corrected image may be dependent on the quality of that correction. Ideally we would like a method of alignment which does not require a separate intensity correction step.

9.2.2 A Solution: Image Partitioning

When evaluating a measure of image alignment, the measure is applied to the whole volume of intersection $\mathbf{x} \in V_M \cap T(V_N)$. In practice, because of the geometry of the MR coil, there will be some points $g(k)$ in the image over which the intensity distortion,

$$g(k) = \{\mathbf{x} | \mathbf{x} \in V_M, G_M(\mathbf{x}) \approx k\}, \tag{9.3}$$

remains constant, or roughly constant. A simple approach is then to evaluate the similarity measure separately over these different sets of points, by forming a partitioning $\mathcal{X} = \wp(V_M)$ of the imaged space.

There are two questions posed by this approach:

- How do we combine our measures of similarity from the separate regions to form a global estimate of alignment for the whole image.
- How do we partition the image so that we can form a useful measure of alignment from each partition.

In order to address the second question, a solution to the first problem will initially be examined.

9.3 Encoding Spatial Location

In order to combine measures of alignment from different regions of the image, it is possible to extend the information theoretic view of alignment to use a partitioning as an extra *channel* of information. The measure of mutual information can be directly extended to multiple sources. In this case we wish to encode the spatial location $\mathbf{x} \in V_M$ of measurement values $m \in M$ in an image in terms of the partition $\mathcal{X} = \wp(V_M)$ in which they occur. We can do this by using a three channel version of mutual information [76],

$$I(M; \mathcal{X}; N) = H(M) + H(\mathcal{X}) + H(N) - H(M, \mathcal{X}, N). \tag{9.4}$$

This relates the entropy of the distorted MR intensities M, their partition $\mathcal{X}$ and the second modality N. In fact, we know that the relationship between the measurements M and their partitioning $\mathcal{X}$ will be fixed. We can therefore use the conditional form,

$$I(M, \mathcal{X}; N) = H(M, \mathcal{X}) + H(N) - H(M, \mathcal{X}, N). \tag{9.5}$$

The relationships between these entropies are illustrated by the Venn diagram in Figure 9.1. The entropies $H(M,\mathcal{X})$, $H(N)$ and $H(M,\mathcal{X},N)$ can be derived from the probability

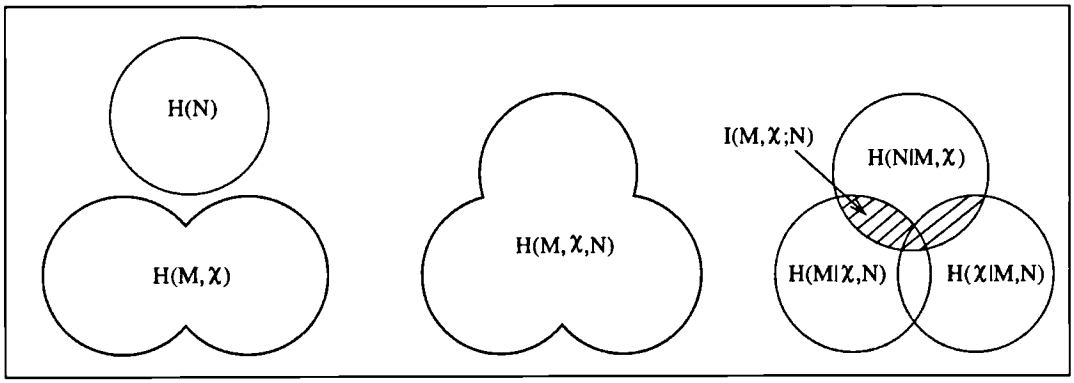


Figure 9.1: A Venn diagram of the entropies involved in the evaluation of three channel partitioned mutual information

of occurrence of measurements M and N, and partitions $\mathcal{X}$,

$$H(M,\mathcal{X}) = \sum_{m\in M}\sum_{x\in\mathcal{X}} p\{m,x\} log\frac{1}{p\{m,x\}} \tag{9.6}$$

$$H(N) = \sum_{n\in N} p\{n\} log\frac{1}{p\{n\}} \tag{9.7}$$

$$H(M,\mathcal{X},N) = \sum_{m\in M}\sum_{x\in\mathcal{X}}\sum_{n\in N} p\{m,x,n\} log\frac{1}{p\{m,x,n\}}. \tag{9.8}$$

Substituting these into equation (9.5),

$$\begin{aligned} I(M,\mathcal{X};N) = \sum_{m\in M}\sum_{x\in\mathcal{X}} p\{m,x\} log\frac{1}{p\{m,x\}} + \\ \sum_{n\in N} p\{n\} log\frac{1}{p\{n\}} - \\ \sum_{m\in M}\sum_{x\in\mathcal{X}}\sum_{n\in N} p\{m,x,n\} log\frac{1}{p\{m,x,n\}}. \end{aligned} \tag{9.9}$$

and given that $p\{m,x\} = \sum_{n\in N} p\{m,x,n\}$ and $p\{n\} = \sum_{m\in M}\sum_{x\in\mathcal{X}} p\{m,x,n\}$ results in an expression:

$$\begin{aligned} I(M,\mathcal{X};N) = \sum_{m\in M}\sum_{x\in\mathcal{X}}\sum_{n\in N} p\{m,x,n\} log\frac{1}{p\{m,x\}} + \\ \sum_{m\in M}\sum_{x\in\mathcal{X}}\sum_{n\in N} p\{m,x,n\} log\frac{1}{p\{n\}} - \\ \sum_{m\in M}\sum_{x\in\mathcal{X}}\sum_{n\in N} p\{m,x,n\} log\frac{1}{p\{m,x,n\}}. \end{aligned} \tag{9.10}$$

which can be rearranged to give a simpler form:

$$I(M,\mathcal{X};N) = \sum_{m\in M}\sum_{x\in\mathcal{X}}\sum_{n\in N} p\{m,x,n\} \log\frac{p\{m,x,n\}}{p\{m,x\}p\{n\}} \tag{9.11}$$

Effectively when evaluating this expression we are forming a three dimensional histogram, where the extra dimension is spatial partition. We are then weighting the influence of the similarities in each partition by the number of measurements in each partition.

Normalised Partitioned Mutual Information

The same approach can be used to develop a normalised version of mutual information (4.11) extended to three sources. As we approach alignment and the shared information between the two images increases, $H(M,\mathcal{X}) + H(N)$ increases with respect to $H(M,\mathcal{X},N)$. From the left of Figure 9.1, we can express this by maximising the ratio,

$$Y(M,\mathcal{X};N) = \frac{H(M,\mathcal{X}) + H(N)}{H(M,\mathcal{X},N)}. \tag{9.12}$$

In chapters 4 and 5 it was shown that the normalised form of mutual information can provide superior response to misalignment. The first set of experiments described in the next section looks at the use of partitioned mutual information $I(M,\mathcal{X};N)$ alone. The experiments of section 9.5 compare the behaviour of both conventional and normalised forms of mutual information.

9.4 Aligning Images with Known Distortion Geometry

Before going on to examine an approach to automatically forming a partitioning, we can first examine an initial application of this partitioned mutual information. Here, because of the nature of the intensity distortion, an appropriate partitioning is fixed and easily defined.

9.4.1 Introduction

At our site MR or CT, and PET images of the pelvis are acquired to aid in the detection and staging of cervical cancer. In this oncology application PET images provide sensitive detection of lesions and nodal involvement. The PET images alone provide very limited anatomical context, whereas MR or CT provides rich anatomical information but often a poor delineation of lesion. The capability to accurately relate locations in the two modalities would provide extremely useful additional clinical information. An example of a manually registered MR and PET image pair was shown in Figure 1.4 in chapter 1.

9.4.2 Aligning MR and PET of the Pelvis

Currently MR and ^{18}FDG PET images are acquired at our site and aligned by the manual identification of corresponding anatomical landmarks. This is aided by the use of an

additional Fluoride tracer which delineates bone structure in the pelvis. This allows the identification of many more bone landmarks which have been found to provide the best registration for the internal pelvic structures [106]. Initially experiments were carried out where a separate F^- tracer image was formed but, due to unpredictable patient behaviour between acquisitions, an accurate alignment of ^{18}FDG using a transformation derived from $^{18}F^-$ could not be guaranteed. As a result a combined ^{18}FDG and $^{18}F^-$ tracer acquisition is made by careful timing of tracer dose. This ensures that the the features delineated by the two tracers are in accurate correspondence.

Although this manual point based alignment only assumes a rigid transformation between the MR and PET acquisitions, it has been shown to provide clinically useful accuracy within the pelvis [106], helping to distinguish regions of bladder, nodal and lesion uptake in PET images.

Approaches to Automated Alignment

In some registration applications it can be possible to use additional information available from the diagnostic image formation process, in order to aid registration of the clinical image with other modalities. In PET imaging of the pelvis, the creation of a diagnostic, attenuation corrected emission PET scan requires two acquisitions, a transmission scan and an emission scan as illustrated in Figure 9.2.

The PET transmission scan is essentially a clinical X-Ray CT scan, although at different energy, and with considerably lower spatial resolution and contrast. As a result it shows bone detail and some soft tissue features. One approach that has been investigated has been to automatically register the MR or CT acquisition to the transmission scan and apply the same transformation to the corrected emission scan. This approach has also been used in manual point based registration. Initial experiments indicated that the increased anatomical structure provided significant improvements in performance. The main problem in using this approach is that the transmission and emission scans are not necessarily in registration, and in fact can often be misregistered by many centimeters. This is often a result of the patient being allowed to leave and re-enter the scanner between acquisitions. In the future it will be possible to acquire both transmission and emission scans simultaneously, but currently the use of this scan in registration does not provide adequate robustness to patient behaviour[1].

The second PET acquisition used in the diagnostic image formation process is the raw emission scan. This is attenuation corrected using the transmission scan as an

[1]It is interesting to note that the often significant misregistration of the emission and transmission scans do not appear to influence the clinical value of the final corrected emission scan. This does not imply that alignment of the emission and transmission scan would not improve final image quality.

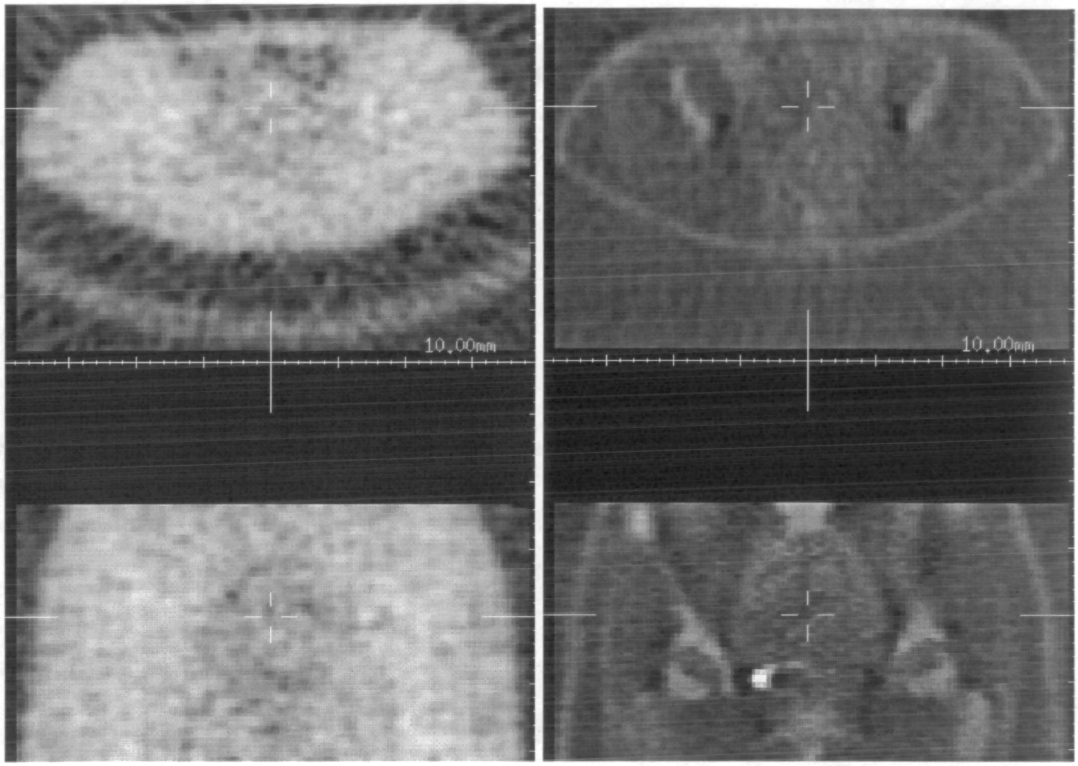


Figure 9.2: Example axial (top) and coronal (bottom) slices through transmission (left) and un-corrected emission (right) scans used for diagnostic image formation.

attenuation map of the patient. The main visible difference between the corrected diagnostic emission and uncorrected emission scan is that the skin surface signal, being less attenuated by tissues, appears much brighter before attenuation correction is applied. This is simply due to the fact that the signal originating from tissue close to the surface of the patient is much less attenuated. This enhanced skin surface may provide additional alignment constraints. Experimentation using similarity measures to align these images with MR images indicated that there was often appreciable misalignment due to deformation of this skin surface between the MR and PET scans.

Overall, although there is an appreciable amount of alternative image data available in the form of uncorrected emission and transmission images, because of the unknown spatial relationship between much of this data, it is only possible to use those structures available in the final clinical ^{18}FDG-Fluoride emission scan for registration.

9.4.3 MR intensity Inhomogeneity

With some MR imaging techniques, notably using a body coil or phased array coils for pelvis imaging, limited extent of the RF field, combined with poor positioning of the coil can lead to considerable intensity distortions in clinical MR images. In these cases the

use of intensity based mutual information as a measure of alignment resulted in visually poor estimates of alignment considerably different from those provided by manual point based registration.

In practice, since the imaging coil is fixed around the patient, the form of the inhomogeneity is constrained. The main loss of field strength (and therefore measurement sensitivity) occurs along the length of the patient in regions falling outside the main imaging volume of the coil (Figure 9.3). The measurement distortion characteristics are then primarily a function of location along the length of the patient. Commonly, in the data sets examined, at least two thirds of the axial field of view is within the imaging region of the coil and therefore provides relatively spatially homogeneous sensitivity. Other measurements falling outside this region have a sensitivity decreasing with distance from the coil. The sensitivity function itself may be empirically modeled from the known design of the coil. The main problem is then the unknown relationship between the coil and the imaging volume which determines the proportion of homogeneous measurements within the imaging region of the coil.

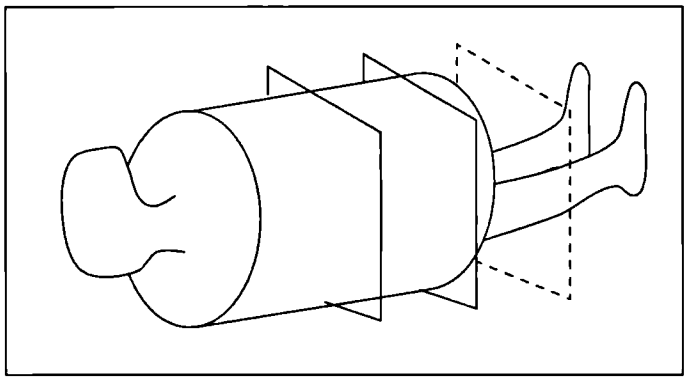


Figure 9.3: Sketch showing relationship between imaging planes, the imaging coil and the patient. Significant distortion occurs in slices falling outside the main imaging volume of the coil.

The secondary component of distortion occurs within plane, again away from the centre of the imaging field. Here though, the the borders of the transaxial slices where most distortion occurs contain mainly zero-signal air around the patient and can therefore be ignored. The in-plane and between-plane distortions may be compared by plotting the intensity at manually traced locations across the image within the same tissue, as shown in Figure 9.4.

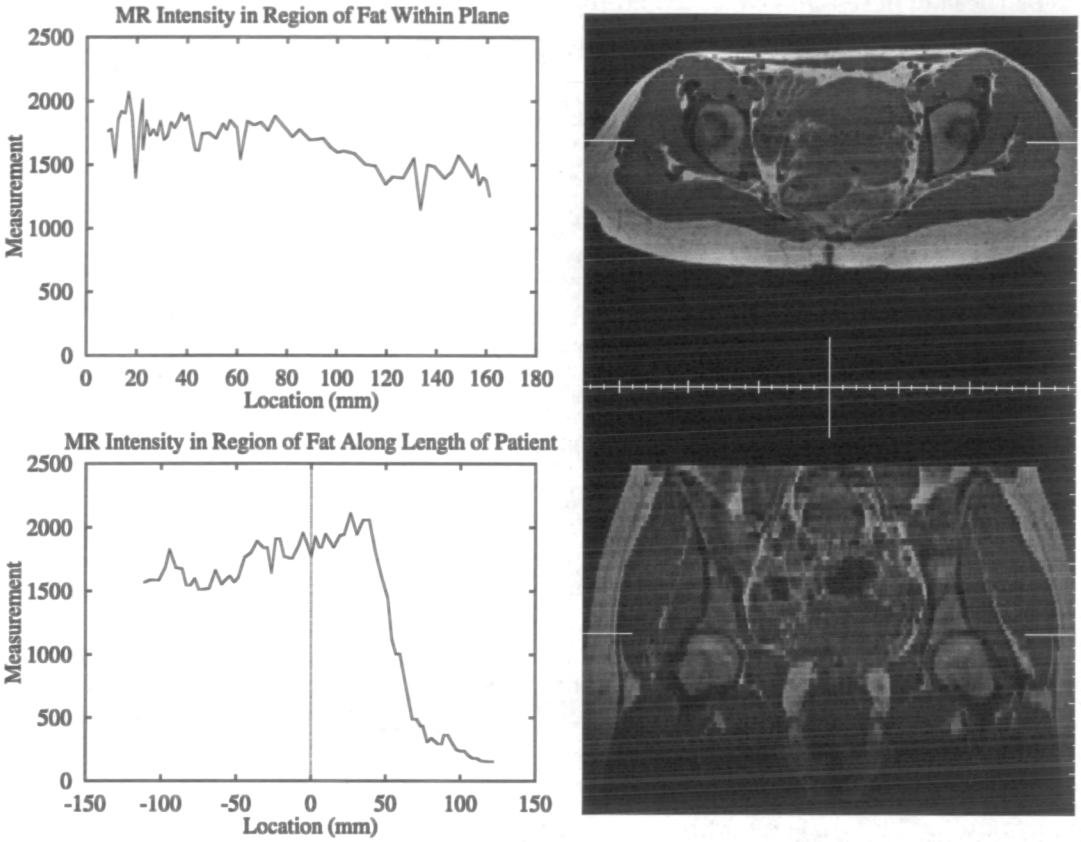


Figure 9.4: Manual traces of points within fat tissue, showing in-plane (top) and between-plane (bottom) distortion in MR intensity.

9.4.4 Evaluation and Optimisation of Image Similarity

The majority of the intensity distortion is in the cranio-caudal axis. By simply encoding MR measurements with slice, it is possible to form partitions in which sensitivity is significantly more uniform. A discrete estimate of the joint probability distribution $p\{m, x, n\}$ is made by simply counting the numbers of pairs of values m and n occurring in each slice separately to form a 3D histogram. By choosing the intensity distorted MR image as the reference coordinate system for measurements, we can directly count the number of measurements m in a given slice with a corresponding interpolated PET measurement n. The partitioning remains fixed with respect to the MR slices during the registration.

As with MR and PET brain image alignment, a multi-resolution approach can be used to increase the efficiency of evaluation during optimisation. Here the partition size is a slice thickness so that at the start of optimisation there are few partitions corresponding to the small number of thick slices. As the optimisation proceeds the image resolution increases and so the number of partitions increases.

For these experiments 64 bins were used to encode intensities in the two modalities. Unlike the brain image alignment experiments, because of the lower spatial resolution of the images, the highest image sampling resolution used to derive a registration estimate was $3mm$.

9.4.5 Experimental Design

Synthetic Data

In order to analyse how the approach can recover a known misregistration we need a pair of images with an accurately known alignment. Because of the nature of the clinical imagery of the pelvis it is difficult to accurately define a ground truth registration. Conventional approaches to estimating alignment using external markers visible in both modalities are affected by movement of the skin surface with respect to the internal organs and bone structure. As a result we have chosen to create a simple synthetic PET image containing bone features simulating those provided by a PET Fluoride tracer scan.

This was achieved by taking an MR image of the pelvis with no significant intensity distortion and hand segmenting regions of the bone structure using an interactive region growing algorithm. This segmented image was then sub-sampled to the same spatial resolution of a PET image and blurred with a Gaussian kernel to simulate a PET point spread function of around $8mm$ FWHM. This image is shown on the right of Figure 9.5.

To simulate the effect of intensity inhomogeneity in MR, an intensity rescaling was applied over the length of the undistorted MR image with the following properties. Given the axial field of view to be $264mm$ divided into 66 slices, if the MR intensity in the original image was given by $\tilde{m}(z)$, the distorted intensity $m(z)$ was given by the following:

$$m(z) = \begin{cases} \tilde{m}(z) & \text{if } z < 132mm \\ \frac{\tilde{m}(z)}{1+k(z-132)^2} & \text{otherwise} \end{cases} \tag{9.13}$$

A value of $k = 0.001$ was used to give an intensity scale factor of 0.054 at the edge of the axial field of view. This was similar to the observed distortion in a number of clinical scans (see for example Figure 9.4). A plot of this function is shown in Figure 9.6 and the resulting MR image is shown on the left of Figure 9.5.

Clinical Data

In order to test the behaviour of the algorithm on actual data, a pair of clinically acquired images were used. The MR image was a conventional T1 weighted 2D acquisition consisting of 66 $4.0mm$ slices of 256×256 $1.716mm$ square pixels. The image included a strong intensity inhomogeneity extending over the lower third of the scan caused by position of the RF coil. The PET image consisted of a combined ^{18}FDG and $^{18}F^-$ tracer

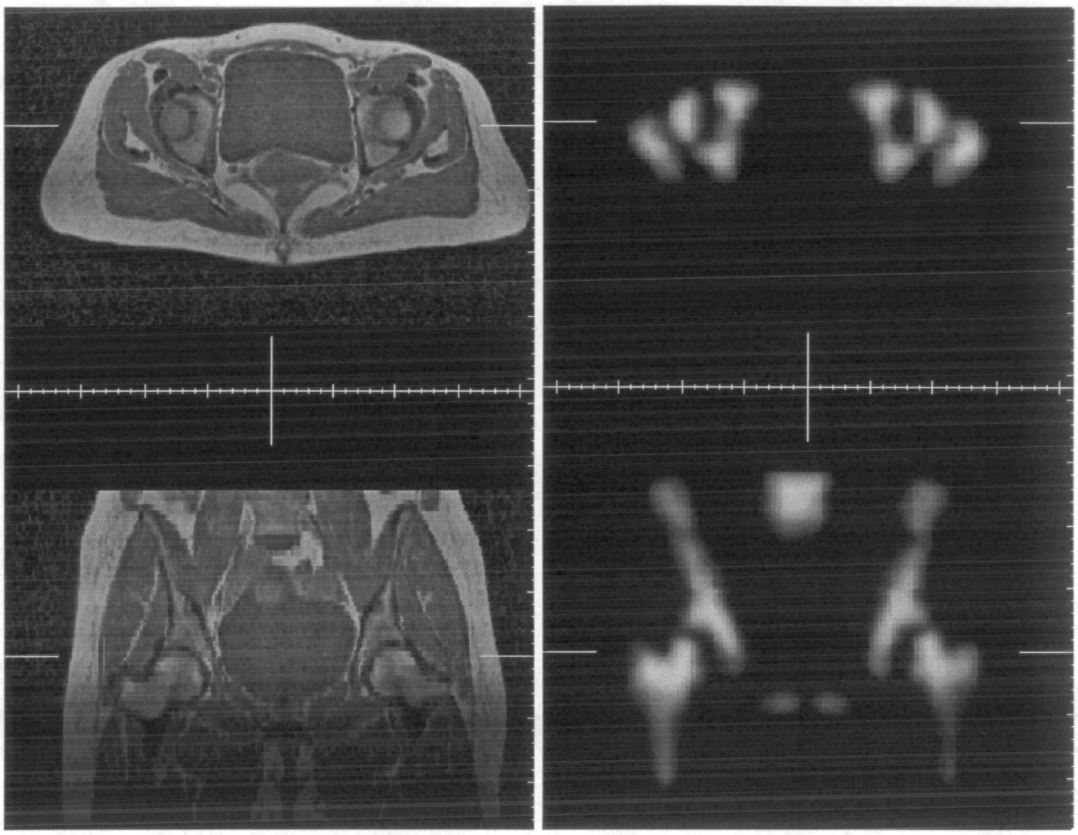


Figure 9.5: Orthogonal slices through the MR image volume with simulated intensity inhomogeneity added (left) and simulated PET bone $^{18}F^{-}$ image (right).

scan showing bone features and and small area of ^{18}FDG uptake. The image volume was made up of 57 3.375mm slices of 128 × 128 3.0mm square pixels. Orthogonal slices through the two image volumes are shown in Figure 9.7.

9.4.6 Results

Synthetic Data

Using the known alignment of the synthetic image pair, values of $I(M;N)$ and $I(M,\mathcal{X};N)$ were calculated for transformations with a displacement in the z-axis around registration. This degree of freedom generally has the largest initial error and is the least well constrained by the structures in the patient. These are plotted as graphs in Figure 9.8. Spatial encoding of intensity has the effect of moving the global optimum to correspond with true alignment. In addition $I(M,\mathcal{X};N)$ appears a much simpler function of misregistration.

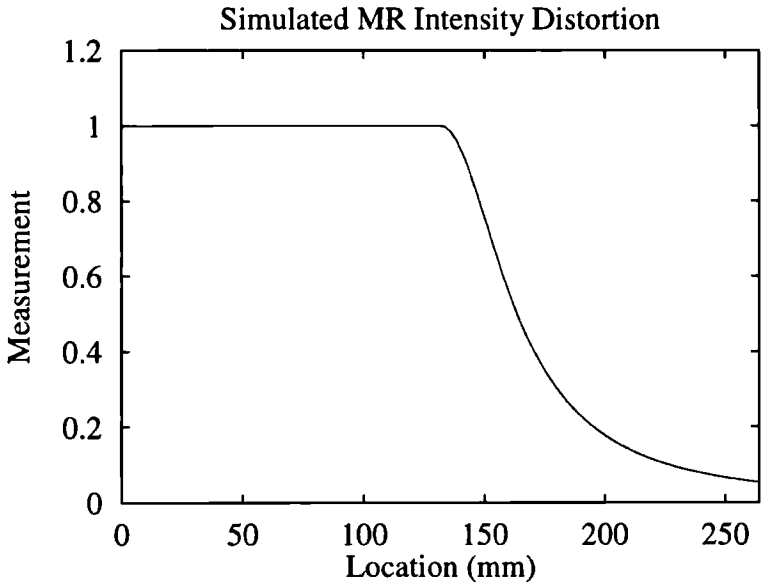


Figure 9.6: Plot of simulated intensity distortion function applied along the length of the patient.

Clinical Data

As with the brain image alignment experiments, an estimate of the centres of the two image volumes aligned and no rotation between the imaging planes was used for the starting estimate for the rigid registration parameters. The final solution was visually inspected using an interactive colour overlay of orthogonal slices of PET intensity onto grey scale MR. The result was found to be visually acceptable, enabling the region of high ^{18}FDG uptake to be anatomically localised as illustrated in Figure 9.9.

A plot of the measure with respect to z axis translation around this final solution in Figure 9.10, shows the difference in the behaviour of the measures for this data. The trace of $I(M, \mathcal{X}; N)$ shows a local optima displaced from the visually acceptable solution which was avoided by the multi-resolution optimisation. This may be caused by the much more complex nature of the real PET image which contains soft tissue features such as the skin surface. As highlighted earlier, the skin surface in PET is commonly distorted with respect to that in MR due to the different design of bed rest. The trace of $I(M; N)$ on the other hand shows a monotonically increasing function of z axis displacement with only a small change in gradient around the visually acceptable registration.

9.4.7 Discussion

The results presented in this section illustrate the feasibility of using an approximate partitioning of the volume to overcome the predominant type of distortion present in MR images of the pelvis. By partitioning intensity with slice, a significant improvement in the

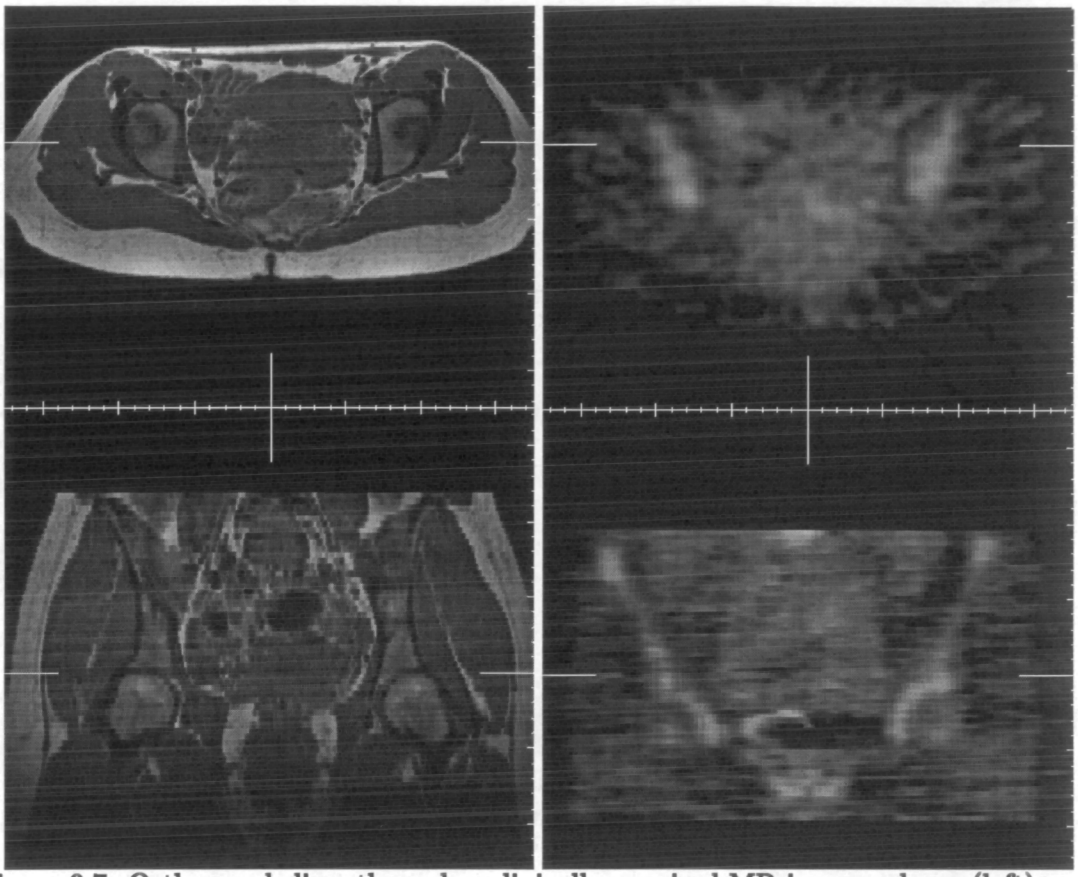


Figure 9.7: Orthogonal slices through a clinically acquired MR image volume (left) and corresponding PET image volume (right), manually registered.

alignment parameter space with respect to PET FDG images is achieved. By employing a simple multi-resolution optimisation scheme, it is possible to achieve a visually acceptable alignment of images for which mutual information derived from intensity alone provided a very poor estimate of alignment.

The problem of fully automated rigid alignment of all types of PET and MR images of the pelvis is still not solved. In particular there are other characteristics of these types of images which mean that the basic assumption for image similarity based alignment will be broken. These will be discussed in more detail in later chapters of the thesis.

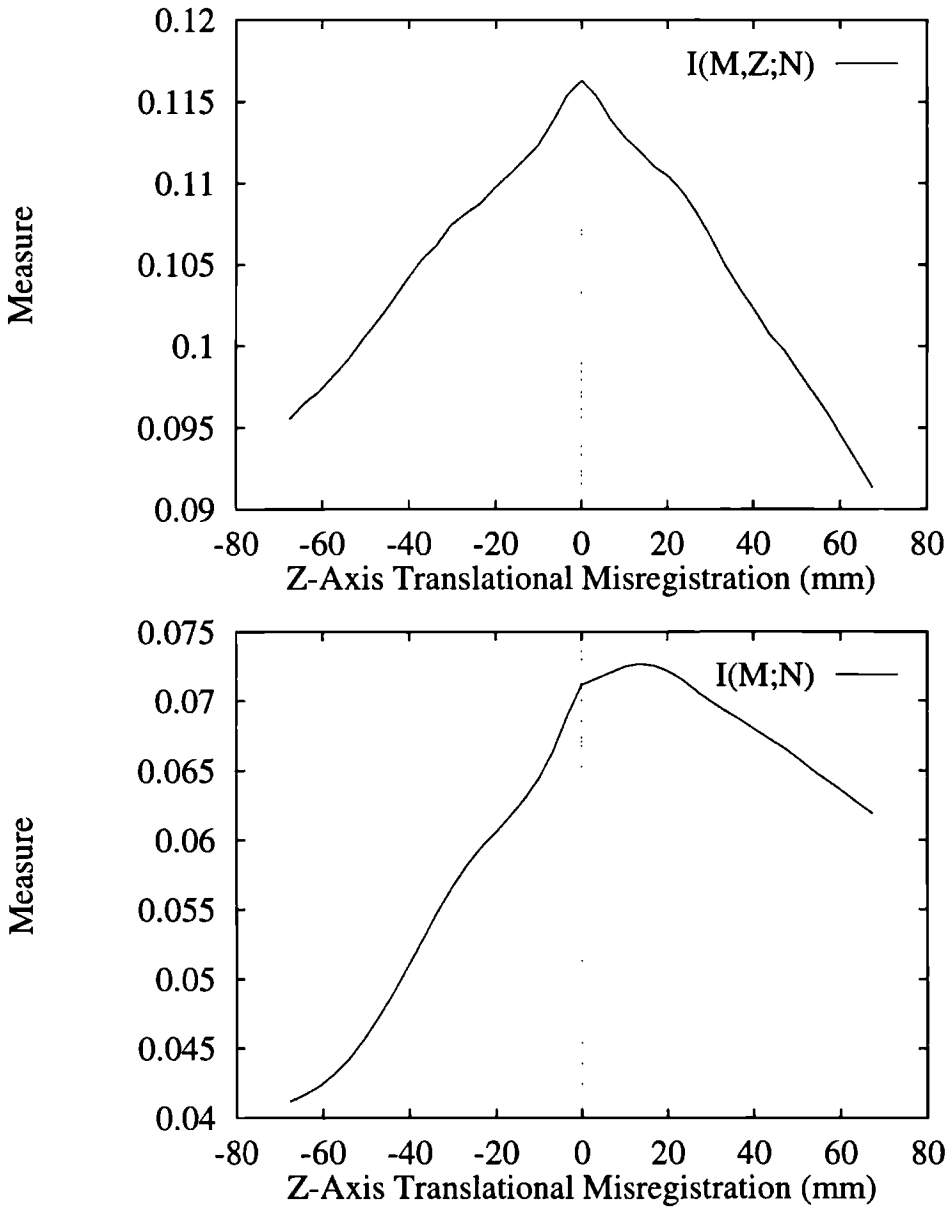


Figure 9.8: Graphs of $I(M;N)$ (bottom) and $I(M,\mathcal{X};N)$ (top) with respect to axial (z) translation of the simulated MR and PET data around the known registration (z=0mm).

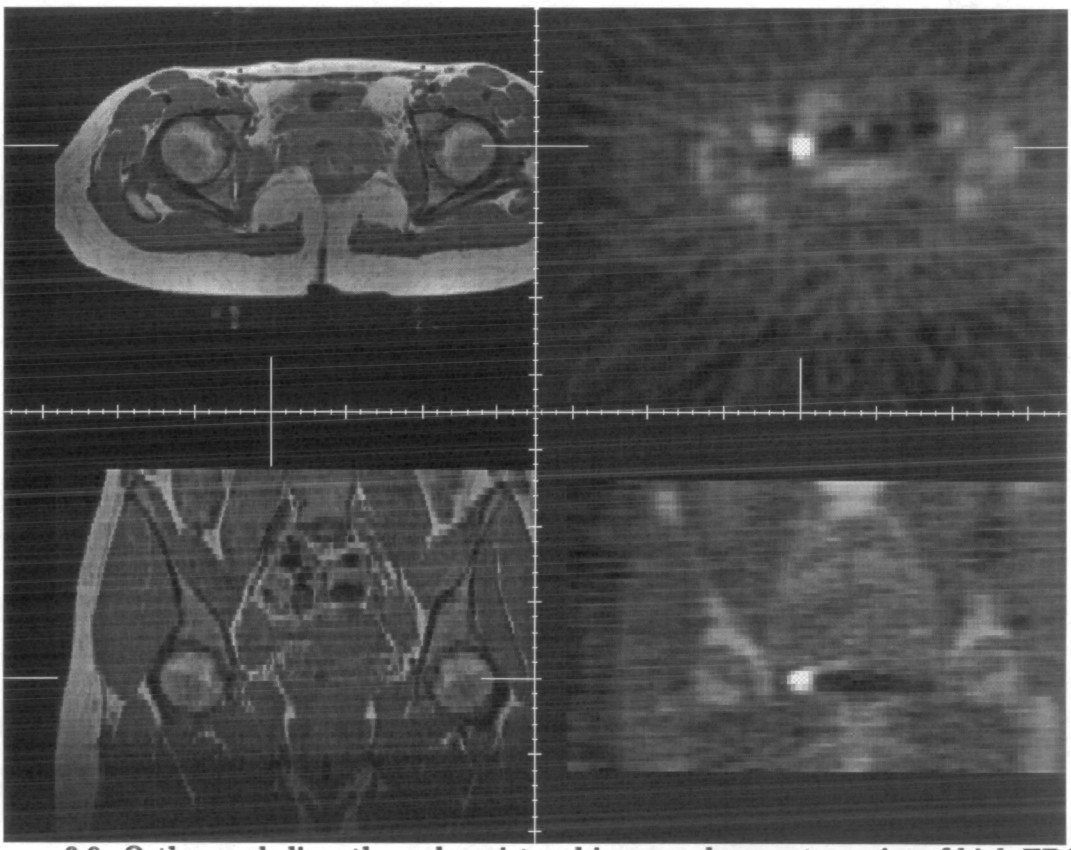


Figure 9.9: Orthogonal slices through registered image volumes at a point of high FDG uptake (right), illustrating its anatomical alignment with MR (left).

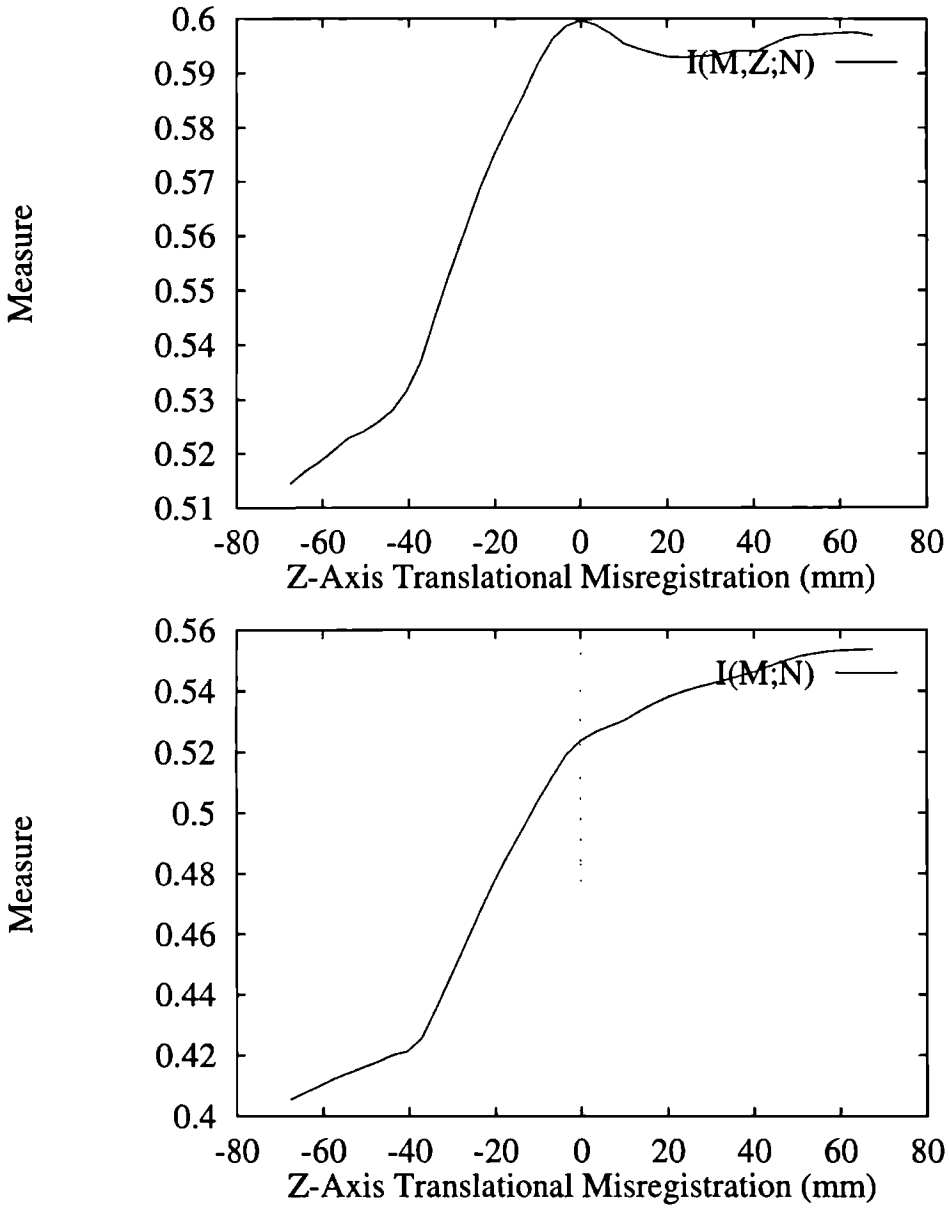


Figure 9.10: Graphs of $I(M;N)$ (bottom) and $I(M,\mathcal{X};N)$ (top) with respect to axial (z) translation of the clinical MR and PET data around the global optimum found by multi-resolution optimisation (z=0mm).

9.5 Automated Partitioning to Recover Unknown Distortion Geometry

9.5.1 A Measure of Partitioning

In the application of the previous section, the orientation of the sensitivity function was fixed relative to the axes of the patient. In some applications this may not be the case. The following view of the problem may be applied to other geometries but, for the moment, we will continue to consider this simple planar partitioning.

If we assume that a pair of images $m(\mathbf{x})$ and $n(\mathbf{x})$ of properties M and N (one intensity distorted and one not) are roughly aligned, and that we have some partitioning of the distorted image $\mathcal{X}$, then we can consider what happens if we vary only the partitioning, rather than the alignment. If a partition contains a uniform region of material present in both images, then this will exhibit a single value (or small range determined by the noise) in the undistorted modality. In the distorted modality each uniform region may exhibit a large range of values.

The probability distribution of measurement values in that region will be determined by the fractional space taken up by each measurement value in that region. Given the underlying region is uniform material, the measurements will be determined by the distortion function $G_M(\mathbf{x})$. If the gradient of $G_M(\mathbf{x})$ has some direction in which it is maximum, for example perpendicular to the plane of a surface coil, and we use a simple planar partitioning in 3D then the range of values exhibited by a region within a partition will be minimised when the partitions are perpendicular to the maximum gradient. As the range of values increases, so the number of voxels with each value will decrease. We can look at this effect in terms of the joint probability distribution $p\{m, x, n\}$ for a given partition as illustrated in Figure 9.11.

So, as with mis-alignment of two modalities, the mis-alignment of partitioning with respect to the direction of maximum distortion results in a dispersion of clusters in the joint histogram. Conversely, as we align the partitioning with the distortion, the spread of values $m \in M$ corresponding to a single value of $n \in N$ falls. This leads to the possibility that we may be able to use information theory to indicate an optimum partitioning.

Taking the case of 3D images, the orientation or direction of a single partitioning axis may be described by a pair of angles, say α and β which defines the partitioning $\mathcal{X}(\alpha, \beta)$. This determines which locations fall within which partitions as illustrated in Figure 9.12. The value of partitioned mutual information is then provided by a function $I(M, \mathcal{X}(\alpha, \beta); N)$, which varies with these angles. It is important that we

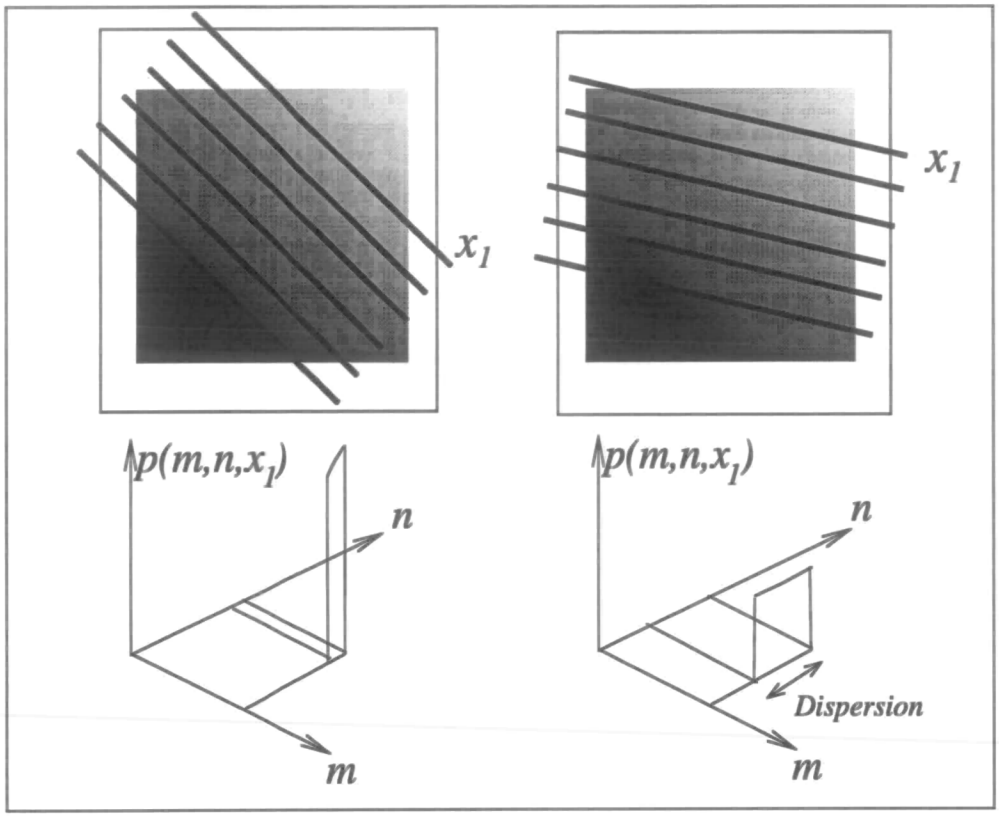


Figure 9.11: As the partitioning of a uniform region of material becomes aligned with the direction of the maximum intensity distortion, so the range of values of the distorted image corresponding to a single value in the other modality is minimised.

still consider the conditional form of three channel mutual information (or normalised mutual information). Subjectively, we don't want to find a partitioning which looks like the distorted image, we want to find a partitioning which makes the distorted image look like the undistorted image.

We can consider what happens to the value of $I(M, \mathcal{X}(\alpha, \beta); N)$ when we vary the partitioning parameters α and β for real data. The actual response will be dependent on the underlying structures and how they are delineated in the two modalities. To simulate the acquisition of a surface coil MR image of the brain, a distortion function similar to that given by equation 9.13 was applied to a head coil image to provide a distortion with a known orientation. Figure 9.13 shows the plots of partitioned mutual information as the orientation of the partitioning angles α and β are varied with respect to the orientation of the synthetic distortion. As can be seen, the measure of partitioned mutual information $I(M, \mathcal{X}(\alpha, \beta); N)$ provides a distinct optimum as the partition becomes aligned with the orientation of the distortion. It also provides, in these axes, a monotonic function of misalignment of partitioning. Information theory can therefore be used to derive a

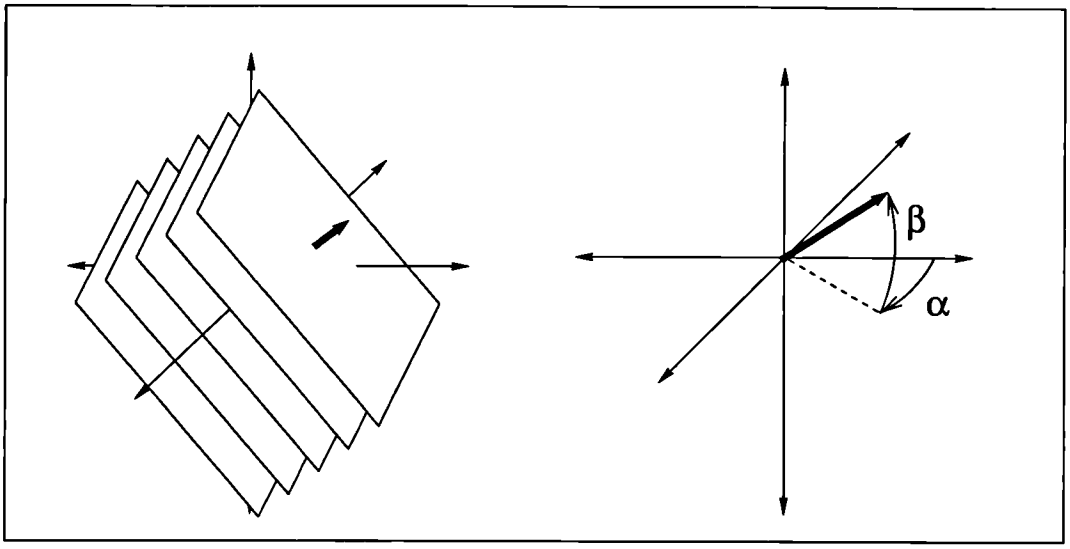


Figure 9.12: Planar partitioning of 3D imaged space determined by two angular variables α and β (note the origin of the rotational axis does not affect the partitioning.)

measure of partitioning as well as image alignment.

9.5.2 Partitioning and Alignment

Given that we can use one measure as both a measure of alignment and partitioning, we can express the overall registration problem as seeking a maximum of $I(T, \alpha, \beta)$. If T is a rigid (six parameter) transformation then we are seeking a maximum with respect to 8 parameters. The difference in response of the measure to changes in partitioning and changes in alignment is a key factor. This will depend on both the nature of the distortion and the underlying structures being imaged and we will investigate this for an important application in brain imaging in the following section.

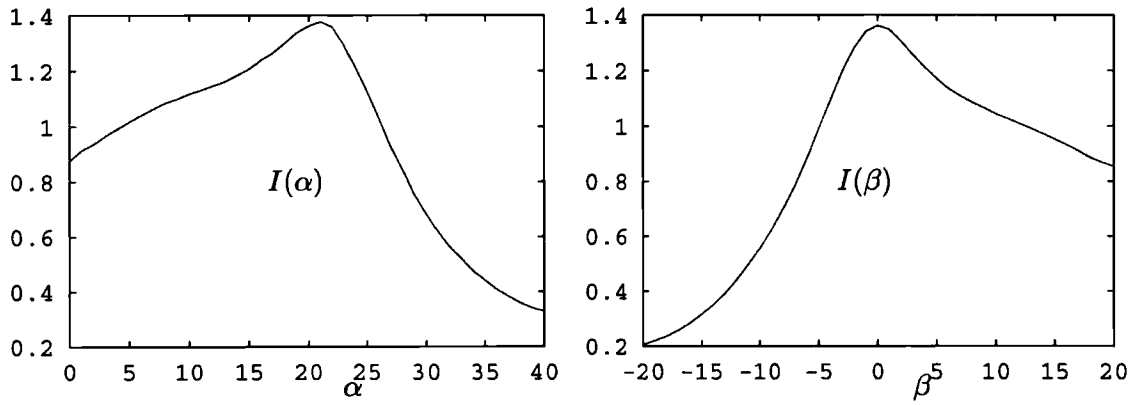


Figure 9.13: For an image of the brain with synthetic distortion applied with orientation $\alpha = 20^o$ and $\beta = 0^o$: Partitioned mutual information evaluated at different partitioning angles α (left) and β (right) (in deg.) between an MR head coil image and an artificially distorted version of it.

9.6 Automated Alignment of MR Surface Coil Angiography of the Brain

9.6.1 Surface Coil MR Imaging

In MR imaging, a surface coil can be used to increase local contrast in a region of interest. Surface coils permit localised imaging with greater signal to noise and efficiency than conventional body and head coils. Surface coils do not however provide uniform sensitivity, and thus are most useful when the region of interest is relatively superficial and focal. The sensitivity profile of surface coils makes them less prone to artifacts from distant structures, while providing images of greater resolution than is possible with larger coils. This is put to use in routine clinical imaging of the spine, temporal-mandibular joints, and shoulders, and specialised applications such as MR angiographic imaging of the brain for surgical planning and guidance [18].

In such cases, sensitivity varies considerably over the imaged field and an appreciable portion of the image can exhibit little or no contrast. This fall off in signal is often used to advantage as a form of data windowing to aid in the Fourier reconstruction of a region of interest by reducing wrap-round of the material at the edge of the field of view. To provide anatomical context it is useful to register these images having local higher contrast and resolution, to a larger volume MR or CT.

9.6.2 Alignment of MR Angiography

Clinically it would be useful to relate T1 weighted anatomical images to the imaged blood vessel structure. The approach used at our site for aligning conventional head coil MR angiography with other modalities, is to register the signal intensity image associated with the angiographic image, to the second modality. This secondary signal intensity contains significant low contrast anatomical structure, and is acquired simultaneously to (and so in registration with) the angiographic measurements, such images are illustrated in Figure 9.14. In the case of surface coil imaging, both the angiographic and signal intensity image contain significant measurement inhomogeneity.

9.6.3 Evaluation and Optimisation of $I(M, \mathcal{X}; N)$

Given corresponding measures $m(\mathbf{x})$ and $n(T(\mathbf{x}))$, we need to decide on how this 'count' contributes to the bins representing the discrete estimate of $p\{m, x, n\}$ say $\hat{p}\{\hat{m}, \hat{x}, \hat{n}\}$. Here $\hat{m}$, $\hat{x}$ and $\hat{n}$ represent the binned values of measurements m, n location x. Tri-linear value interpolation [92] is used to form an estimate of corresponding measurement $n(T(\mathbf{x}))$, as described in section 5.3.3. These values $\{m, n\}$ are then mapped directly to the nearest bins $\{\hat{m}, \hat{n}\}$.

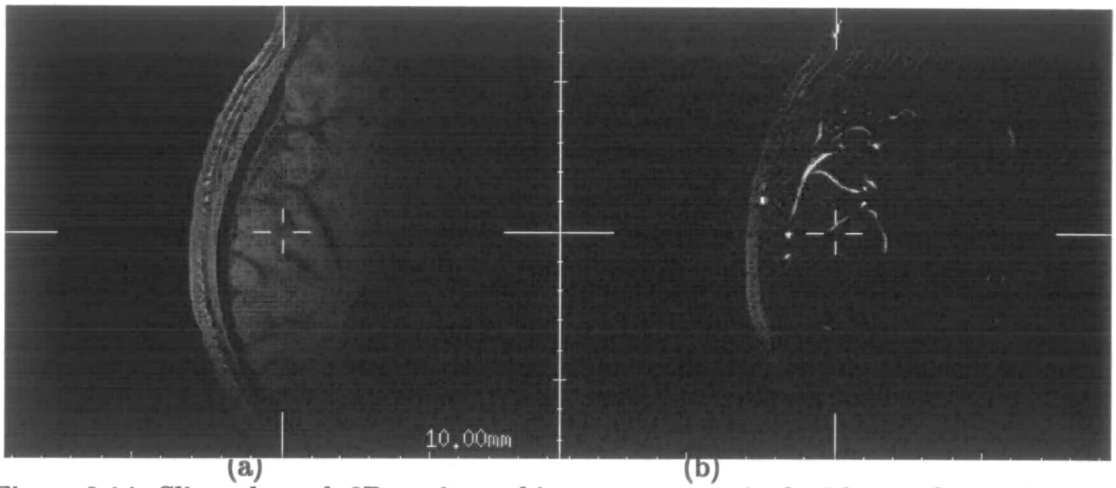


Figure 9.14: Slices through 3D angiographic sequence acquired with a surface coil showing modulus flow image (left) and secondary signal intensity image (right) used for registration with other modalities

Unlike the application in the previous chapter, we must also consider the relationship between the discrete partitioning planes and the discrete sampling of the distorted image. We can treat the partitioning parameters as a transformation $\tau(\alpha, \beta)$ from the discrete voxel locations in the distorted image to a discrete partition space. In order to ensure there is a smooth response of $p\{m, x, n\}$ as the parameters α and β are varied we must employ a form of interpolation. In this case linear volume interpolation was used. This is an adaption of an approach proposed by Maes *et al.* [63, 64] who employed *partial volume* interpolation as an approach to estimating contributions to the joint probability distribution of image *values*. Experimentation with this technique has indicated some limitations, so here we use conventional linear value interpolation to estimate intermediate *values*.

The approach though is suited to evaluate contributions to the joint histogram at intermediate spatial partitions. Given a location $\mathbf{x}$ in distorted image $m(\mathbf{x})$ will fall between two nearest partitions say $\hat{x}_1, \hat{x}_2 \in \mathcal{X}$, we can evaluate the distance d_1 and d_2 of the true location $\tau(\mathbf{x})$ to these partitions as a fraction of a partition, so that $d_1 + d_2 = 1$. A fractional contribution is then made to the two nearest partition bins,

$$\delta p\{\hat{m}, \hat{x_1}, \hat{n}\} = d_2/Q \tag{9.14}$$

and

$$\delta p\{\hat{m}, \hat{x_2}, \hat{n}\} = d_1/Q, \tag{9.15}$$

where Q is the total number of points in the volume of overlap.

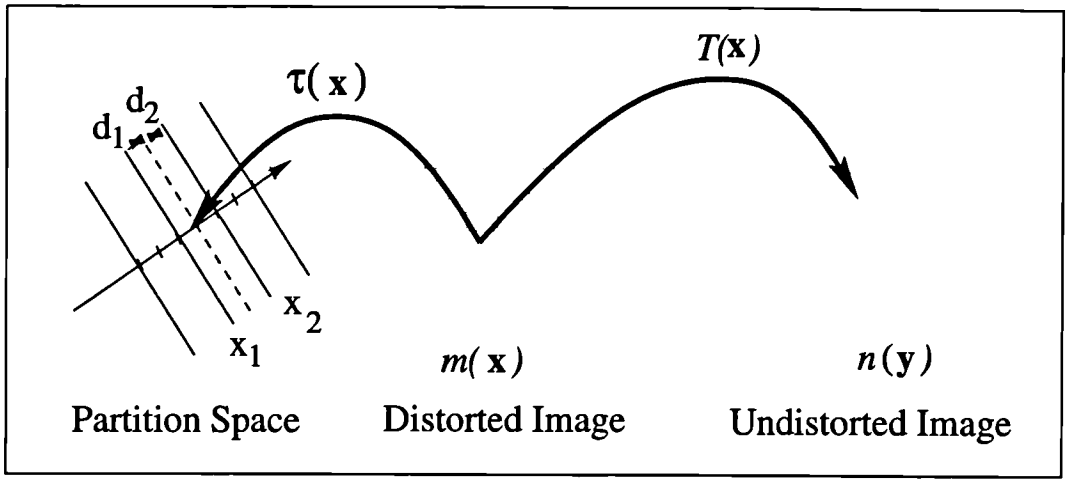


Figure 9.15: Evaluation of contributions to the joint histogram requires two continuous transformations between discrete spaces, Each requiring a form of interpolation: One $T(x)$ between modalities (Tri-Linear Value Interpolation), and one between the distorted image and the planar partitioning (Bi-Linear Volume Interpolation).

Sampling

As described in section 5.3.2, the images are first resampled to $1.5mm$ cubic voxels using a Gaussian kernel to reduce resolution. Multi-resolution versions of the images are then created using a Gaussian kernel. For a given transformation T a discrete estimate of $p\{m, n\}$ is evaluated by forming a histogram where measurements in the two modalities are mapped to 64 intensity bins. In this test image a single axis was used for measurement encoding, approximately aligned with the axis of the surface coil (left to right in the axial slice of Figure 9.16). Although the inhomogeneity is a function of the three axes, initial experimentation indicated that using only this single axis provided satisfactory results. MR measurements are thus divided into location bins, each corresponding to image sampling points, their size being determined by the image resolution during the multi-resolution optimisation.

9.6.4 Experimental Design

MR-MR Data

In order to validate the final registration estimate, a pair of images were acquired of a volunteer using first a surface and then head coil for imaging. Using a 1.5T MR scanner (Philips ACS) we obtained T1/proton density weighted images of a volunteer which covered the whole head using the head coil, and the frontal region using an 8.5cm surface coil. Identical 3D gradient echo scans were made of the whole head with imaging parameters: 5.7/13 msec (TE/TR), flip angle of 12 degrees, and 22cm field of view

for voxel dimensions of $0.858 \times 0.859 \times 1.0mm$. The images were acquired using the same parameters and with care to ensure there was no motion of the patient between acquisitions. The images were therefore assumed to be spatially aligned. Corresponding transaxial and sagittal slices through these images are shown in Figure 9.16.

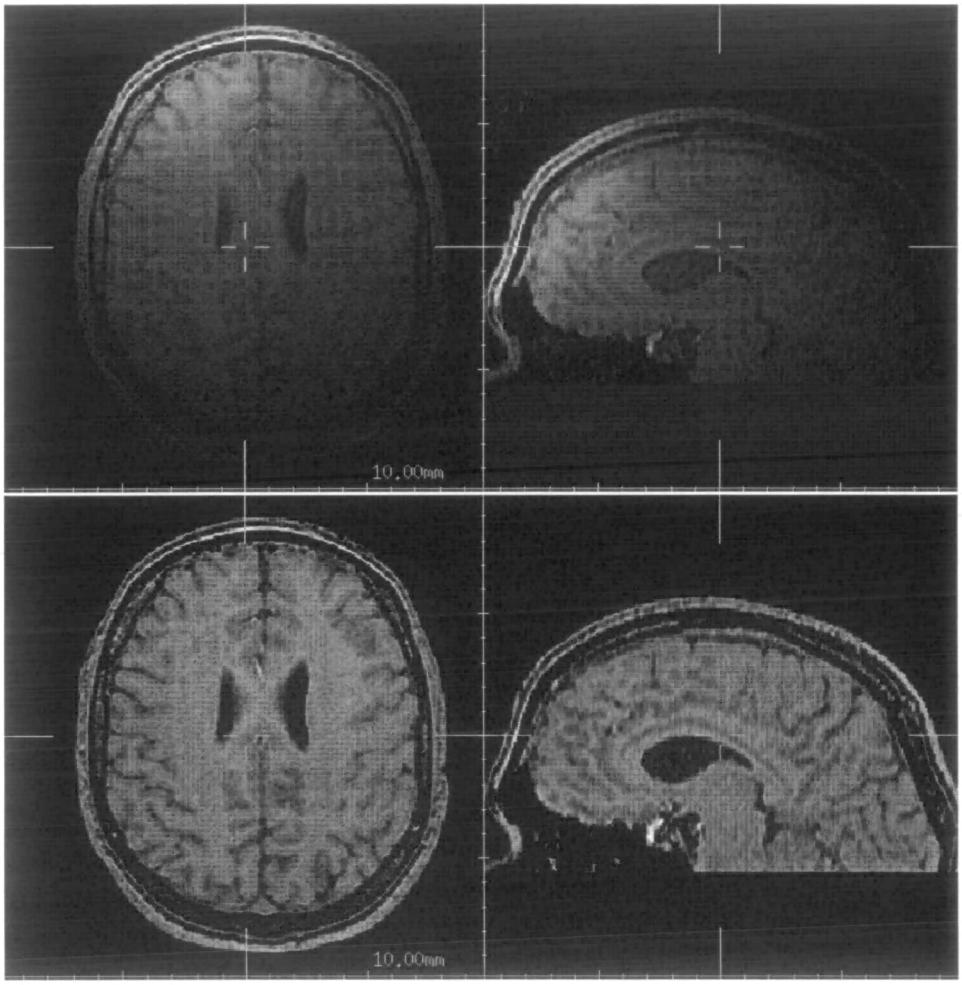


Figure 9.16: Slices through 3D test images acquired with a surface coil (top) and head coil (bottom), illustrating differences in sensitivity across transaxial slices (left) and sagittal slices (right).

MR-CT Data

A second set of data was acquired to examine the alignment of surface coil MR with CT of the head. Here, in order to provide an independent registration estimate an MR-T1 CT image pair from the Vanderbilt retrospective registration project [103] was used. This image data is described in more detail in section 6.3.1. An additional surface coil image of a bottle containing a uniform solution of $770mg$ of $CuSO_4$ per litre of water was acquired locally to form a surface coil distortion map. This is illustrated in Figure

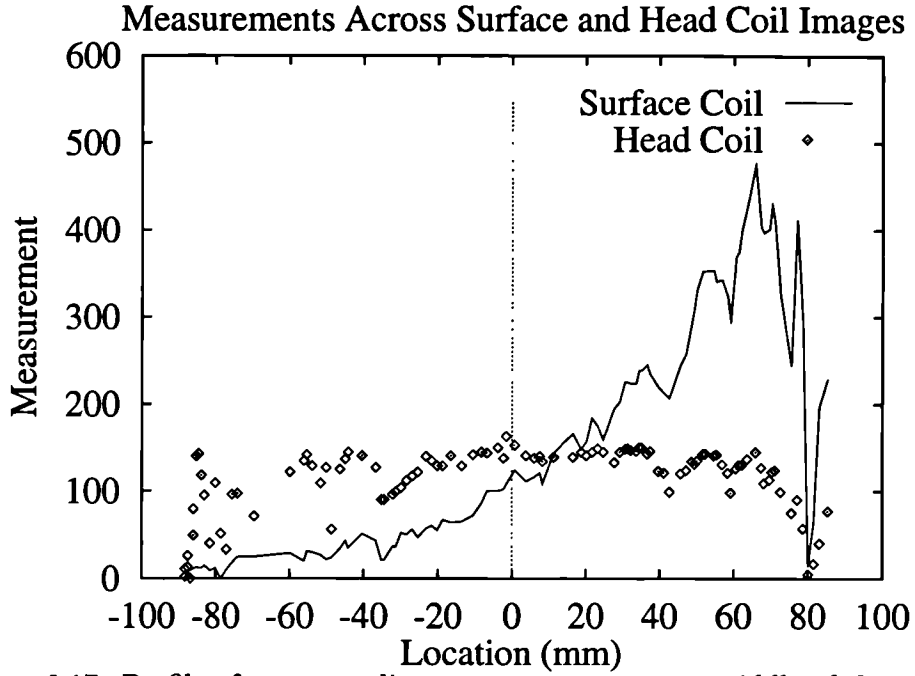


Figure 9.17: Profile of corresponding measurements across middle of these slices, from left to right in the axial plane.

9.18. This was then applied as an intensity scaling factor to the Vanderbilt MR image with the maximum value in the bottle image corresponding to a scaling factor of 1. This provided a simulated surface coil image (Figure 9.19).

Starting Estimates

In order to investigate the parameter space provided by conventional and spatially encoded measures we have taken the known alignment of the test image pairs and perturbed the estimate by a random translation of size 20mm and a random rotation of size 20 degrees, distributed as described in section 6.6.1. Fifty of these random orientations were then used as starting estimates to the multi-resolution optimisation scheme. The final estimates were then recorded to give an indication of the precision and accuracy of estimates.

9.6.5 Results

MR(surface coil)-MR(head coil)

The mean and standard deviation of the displacement of the six alignment parameters from the correct solution for the 50 estimates are shown in Table 9.1 for both the mutual information and normalised mutual information measures. When using only intensity values to derive the measure of alignment, the mean estimate provided by $I(M;N)$ is

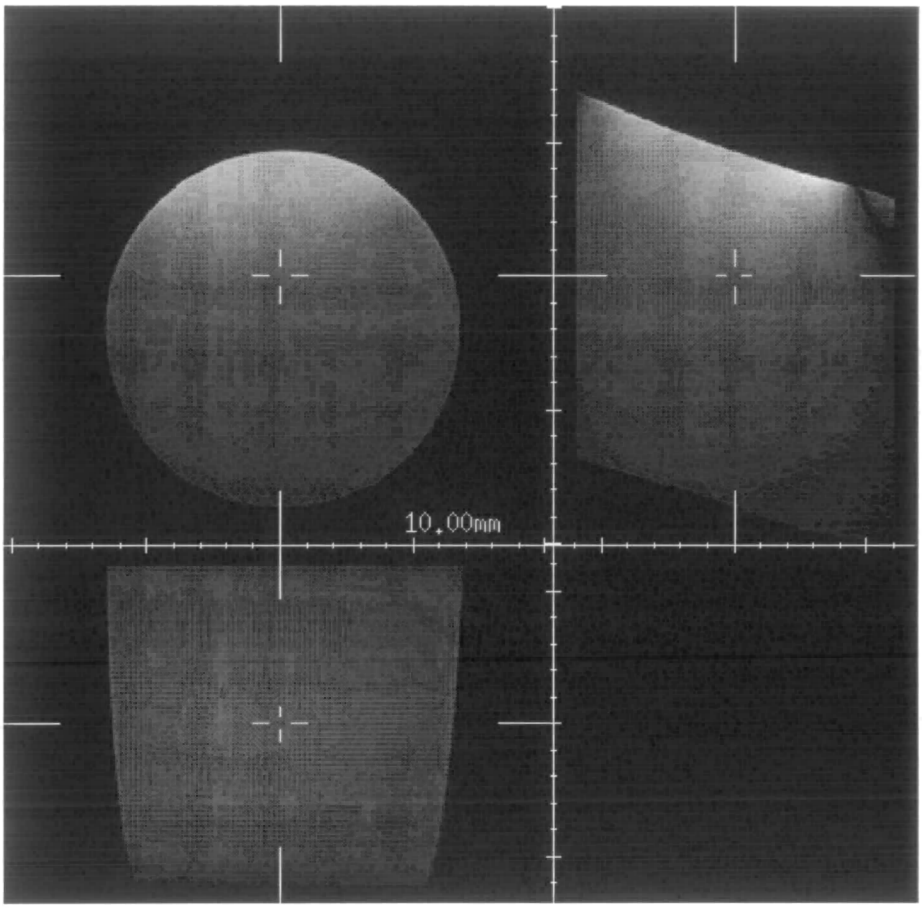


Figure 9.18: Orthogonal slices through intensity distortion 'map' of surface coil produced by imaging a bottle containing uniform solution.

shifted significantly from the true solution, in addition there is a large spread of results around the mean. Normalised mutual information $Y(M;N)$ is affected even more, with larger standard deviations for many parameters. These spread of estimates are illustrated by the two lower scatter plots in Figure 9.20 which show the error in the translation and rotation components of each estimate. In particular this shows that normalised mutual information $Y(M;N)$ provides two distinct clusters of estimates, while mutual information $I(M;N)$ provides a more random spread.

When using partitioning, the mean estimates provided by both measures ($I(M,\mathcal{X},M)$ and $Y(M,\mathcal{X},M)$) are very close to the gold standard and the spread of estimates is significantly reduced. For both measures, there is a slight shift of the mean estimate away from alignment in the y axis which is perpendicular (vertical in the transaxial slice in Figure 9.16) to the coil axis orientation, but this is of a size less than the image voxel dimensions. This may be a result of slight movement of the volunteer between scans or calibration of the MR scanner.

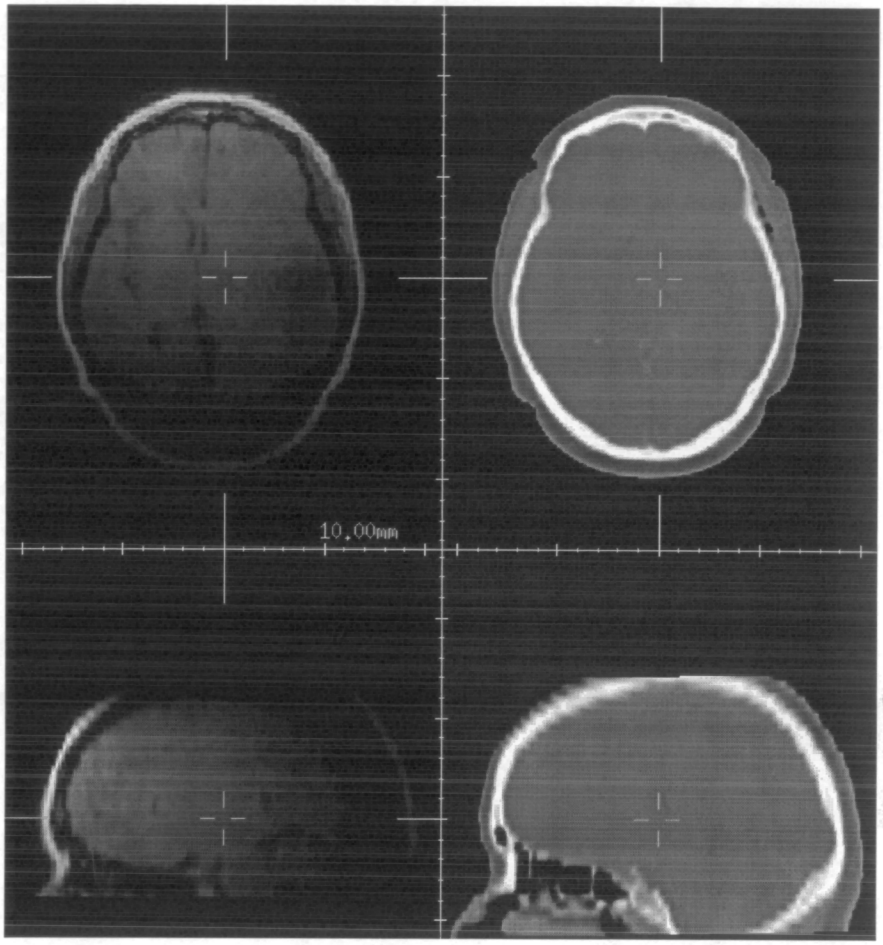


Figure 9.19: Transaxial and sagittal slices through artificially distorted MR image of the brain and corresponding CT image from the Vanderbilt registration database.

Simulated MR(surface coil)-CT

As with the previous experiment statistics of the results using both measures are tabulated (table 9.2) and shown graphically as a scatter plot Figure 9.21). Here, unlike MR-MR alignment, both conventional intensity derived measures ($Y(M;N)$ and $I(M;N)$) provide only one main cluster but shifted significantly from the correct solution. In both cases the solution includes a significant shift in the y axis away from the marker based solution. This corresponds approximately to a displacement of the head coil MR away from the surface coil along the axis of the surface coil.

The use of a partitioned evaluation of the measures provides significantly improved clustering of estimates, with a mean estimate close to the marker solution. The mean estimates provided by the two measures are different but are both within the expected accuracy of the marker based estimate and so we cannot say which is closer to the correct solution. It is interesting to note that although the range of registration parameter

MR-MR

	Measure			
Parameter	$I(M;N)$	$Y(M;N)$	$I(M,\mathcal{X};N)$	$Y(M,\mathcal{X};N)$
$t_x(mm)$	0.01(0.31)	-1.03(4.16)	0.17(0.05)	0.19(0.072)
$t_y(mm)$	-0.44(0.44)	2.40(1.34)	0.13(0.03)	0.15(0.023)
$t_z(mm)$	-1.47(0.90)	14.18(8.67)	-0.01(0.01)	-0.03(0.001)
θ_x°	-0.01(0.31)	-1.48(1.98)	-0.05(0.02)	-0.04(0.021)
θ_y°	-0.35(0.45)	-2.82(10.18)	-0.11(0.04)	-0.12(0.059)
θ_z°	0.16(0.54)	-0.176(3.25)	0.05(0.05)	0.03(0.027)
α°	-	-	70.8(0.03)	71.0(0.17)
β°	-	-	-30.2(0.05)	-31.1(0.31)

Table 9.1: Mean (Standard Deviation) of errors in rigid registration parameters and estimates of partition angles (α, β), from 50 random starts estimated by mutual information and normalised mutual information derived from measurements only, and from measurements encoded with spatial location.

estimates are comparable using the conventional and normalised mutual information, the partitioning parameters appear less well defined for the normalised form $Y(M,\mathcal{X},N)$ (See Table 9.2).

Simulated MR-CT

	Measure			
Parameter	$I(M;N)$	$Y(M;N)$	$I(M,\mathcal{X};N)$	$Y(M,\mathcal{X};N)$
$t_x(mm)$	0.49(0.50)	0.68(0.19)	0.02 (0.08)	0.09(0.12)
$t_y(mm)$	2.72(2.17)	3.61(0.04)	0.07(0.08)	-0.03(0.06)
$t_z(mm)$	-0.38(1.98)	0.64(0.07)	0.32(0.04)	0.34(0.05)
θ_x°	-1.91(7.40)	0.90(0.08)	-0.08(0.16)	-0.13(0.15)
θ_y°	-0.43(0.87)	-0.51(0.36)	0.03(0.27)	-0.16(0.26)
θ_z°	-0.85(1.18)	-0.61(0.24)	0.04(0.14)	-0.12(0.16)
α°	-	-	16.1(0.75)	16.0(1.0)
β°	-	-	-110.7(0.60)	-110.6(1.2)

Table 9.2: Mean (Standard Deviation) of errors in rigid registration parameters and estimates of partition angles (α, β), from 50 random starts estimated by mutual information and normalised mutual information derived from measurements only, and from measurements encoded with spatial location.

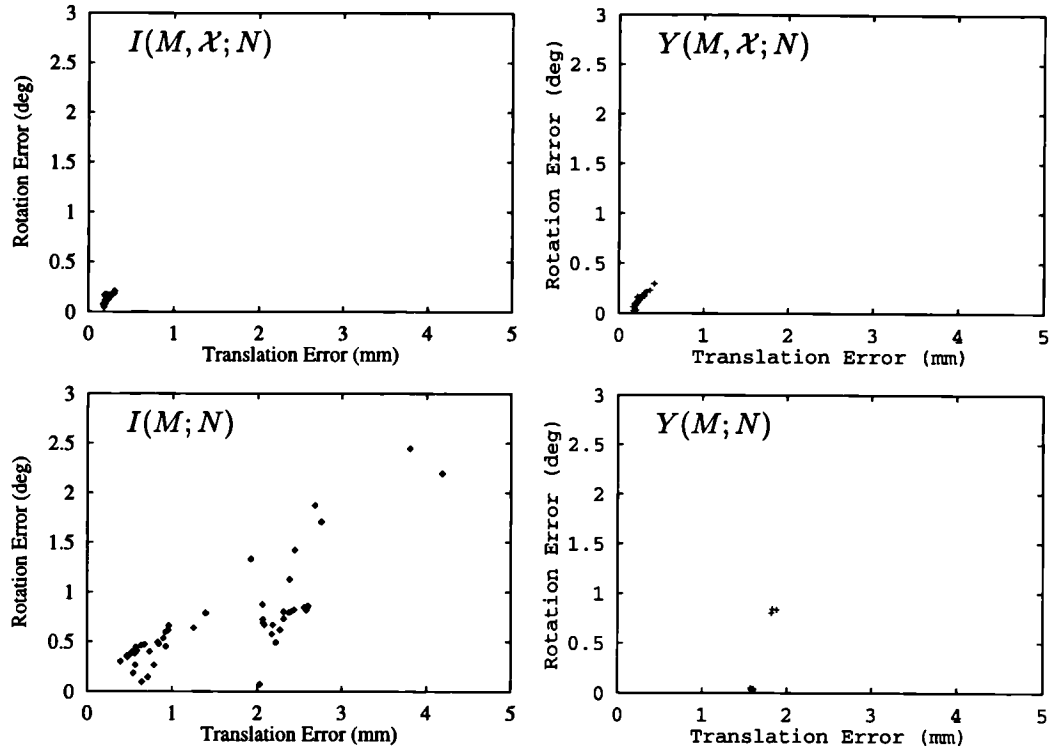


Figure 9.20: MR(Surface Coil)-MR(Head Coil): Lengths of translation and rotation error vectors of parameters estimated by optimisation of different measures.

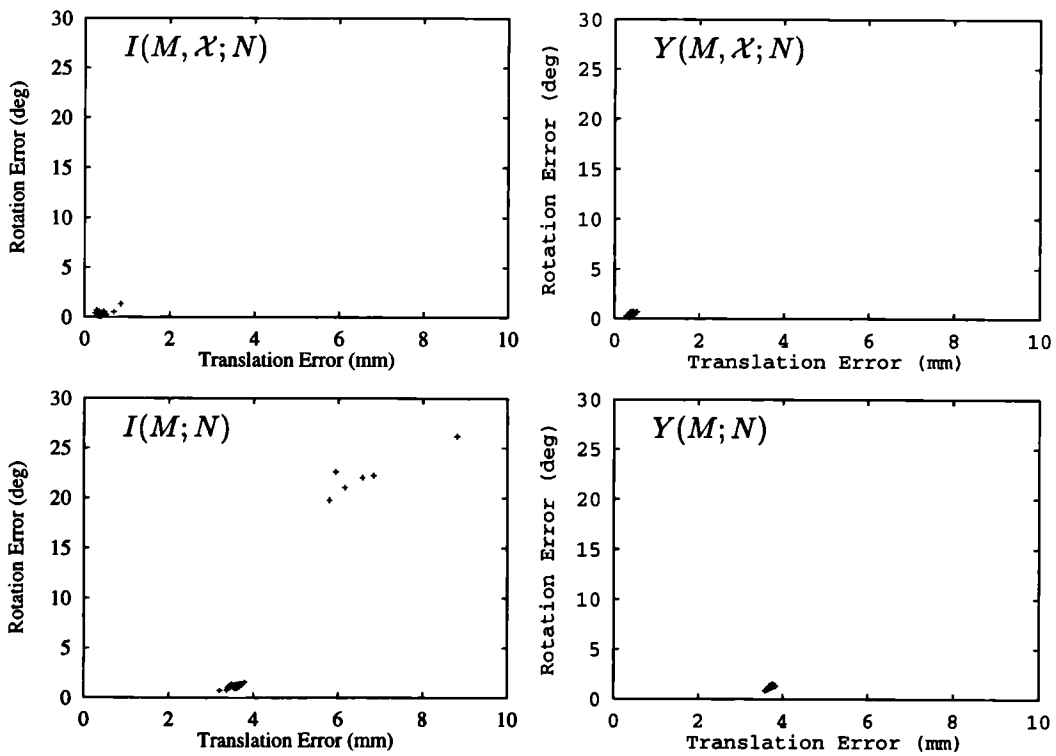


Figure 9.21: MR(Surface Coil)-CT Lengths of translation and rotation error vectors of parameters estimated by optimisation of different measures.

9.7 Summary

In some types of 3D medical images, severe intensity distortion prevents the use of conventional intensity based measures of alignment. This chapter has been concerned with the use of a spatial partitioning of intensity values in deriving entropy based measures of alignment. In cases where the geometry of the distortion with respect to the patient is known, for example in MR-PET pelvis image alignment, a simple planar partitioning provides a significant improvement in the registration parameter space.

By using the registration measure as an indication of the quality of image partitioning it is possible to address the more general problem of surface coil image alignment, where the orientation of the distortion with respect to the patient is unknown. By using a multi-resolution optimisation of both alignment and partitioning parameters, a completely automated approach is feasible. The results of applying this to aligning surface coil MR with MR head coil or CT images of the same scene show the possibility of completely automated and accurate recovery of alignment, where conventional approaches fail.

Further work is necessary to examine the response over a larger range of examples of surface coil locations with respect to the patient and to more clinical examples of normal and abnormal patient anatomy.

Chapter 10

Region Topology and Image Similarity

10.1 Introduction

In many medical images, particularly those of the brain, the fundamental assumption of our approach, that the co-occurrence of the most probable values in the two modalities is maximised at registration, is valid. In the previous chapter, an extension of entropy based measures of alignment was proposed to enable the registration of images where this assumption did not apply. In that case the assumption was invalid because of distortion in measurement values introduced by the imaging process. This chapter examines approaches to dealing with more fundamental cases where the content of the scenes themselves invalidate the assumption.

This chapter is somewhat different from many of the others in that it discusses approaches and work in progress, rather than algorithm design or validation. It does not provide complete or validated approaches to medical applications.

10.2 When Voxel Similarity Fails: Why we still Need High Level Vision

An example of a scene in which a simple measure of image similarity cannot provide an indication of alignment is illustrated by the sketch in Figure 10.1. Here there is a simple binary image of a rectangular box in the two modalities, while within the rectangular box either a circle or a square is delineated. The circle and square do not correspond spatially, but the box around them does. The problem here is that the misaligned regions of the square and circle at registration, are greater in area than those parts of the misaligned box when the circle and square are roughly aligned. Similarity measures such as those derived from entropy will roughly align the square and circle, rather than the smaller area of the rectangular box around them.

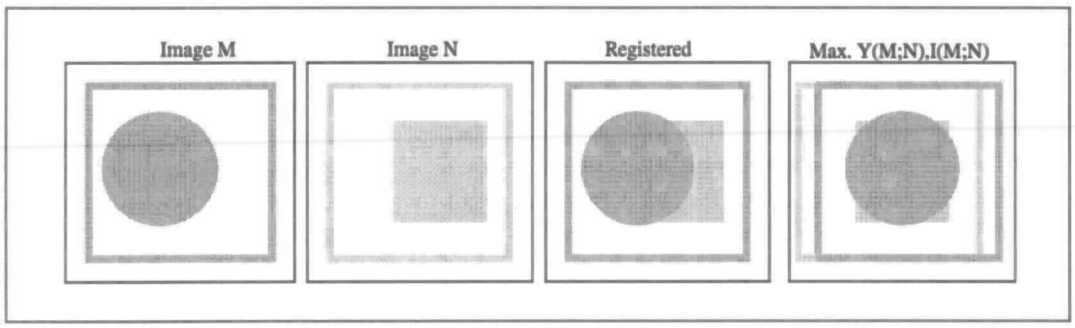


Figure 10.1: A sketch of views of a scene in which voxel similarity measures will not provide a maximum at registration.

Without some higher level knowledge about the correspondence of regions within a scene it is not possible to define a simple similarity based measure of alignment. In fact, because the extent of the rectangle perpendicular to the left-right translation is greater than the extent of the cube and circle there will be a local optima at 'registration'. It is though possible to construct scenes where there will be no local optima at registration.

There are examples of views provided by different modalities where the intensity similarity assumption is violated, but where additional spatial information may be used to allow their alignment, without the need for a direct knowledge of region correspondence. The following sections described such a case and how it is possible through simple assumptions about the topology of regions within the images to provide a modified measure of alignment.

10.3 Using Region Topology in Image Registration

10.3.1 Connected Component Labelling

It is rare for a medical image volume to contain large areas of non corresponding values. It is though possible that the same value may be exhibited by unconnected regions of tissue in the field of view of both modalities. This poses a problem to image similarity when only some of these unconnected regions correspond. This may be because, for example a tracer is used in one modality which detects the same activity in tissue exhibiting different properties in the other modality.

A simple modification of the earlier model illustrates this case in Figure 10.2. Here, as in the previous figure, maximum overlap of values does not correspond to our definition of alignment.

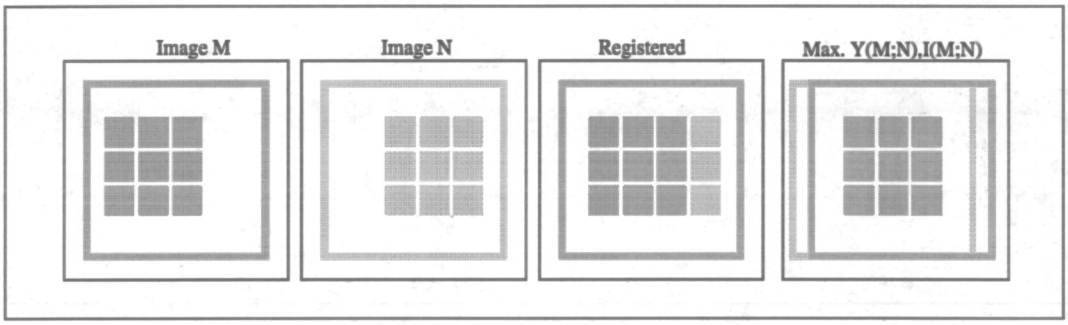


Figure 10.2: A sketch of views of a scene provided by different modalities in which voxel similarity measures will not provide a maximum at registration.

Here we may be able to assume that all corresponding regions within the two modalities are spatially connected. If we apply a simple connected component labelling of the image voxels to create a label image in which all unconnected regions exhibiting the same value are labelled with different values, the response of a similarity measure is modified significantly. In this case the value exhibited by the squares will be partitioned and the proportion of the image occupied by each label will be much smaller, as illustrated in Figure 10.3. As a result, the transformation at which one image best explains the values in the second now corresponds to true alignment. So, without introducing higher level information about region correspondence, we have provided a useful measure of image alignment.

10.3.2 Using Connected Component Information in Image Alignment

In this simple example we have looked at the problem in terms of the content of the scene being imaged. In practice we manipulate images of a scene which are limited by

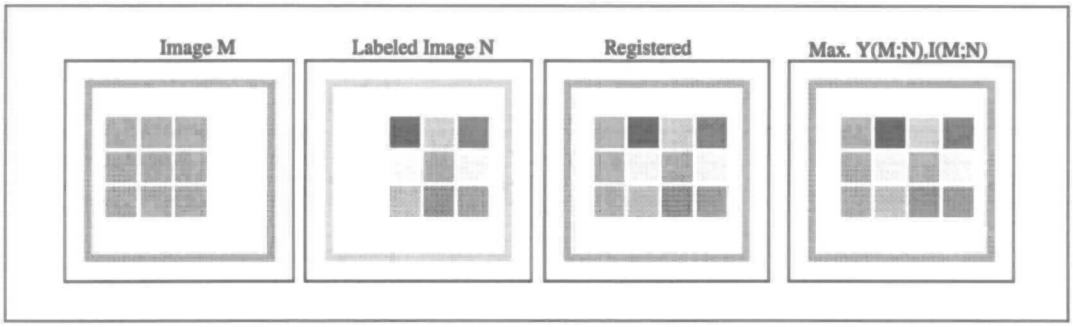


Figure 10.3: A sketch of views of a scene from Figure 10.2, where a connected component labelling (here represented by random grey values) has been applied to one image to distinguish regions which exhibit the same value but are not spatially connected.

spatial resolution and noise. The distribution of values exhibited by materials in the scene are continuous and it is difficult to define an accurate labelling of an image. In effect we have returned to the problem of corresponding object segmentation in different modalities described in chapter 2. We cannot easily define a threshold to delineate corresponding regions in two modalities. In this section we will examine an alternative approach where we do not derive our estimate of alignment simply from a segmented image.

In practice, for an image of a scene such as that illustrated in Figure 10.3, we may be able to create a rough labelling of the image which distinguishes between unconnected regions exhibiting the same value. The approach taken in this work is to use this labelling as an additional source of information *explaining* the original image intensities. This is achieved using the same approach as was used in chapter 9, by introducing the labelling as a third source of information in the registration measure. Given we have a set of labels $L = \{l_0, l_1, \ldots l_i\}$, about image $m(\mathbf{x})$ with values $m \in M$, which we wish to align with image $n(\mathbf{y})$ with values $n \in N$, then we may employ equation 9.5 so that we search for the maximum of,

$$I(M, L; N) = H(M, L) + H(N) - H(M, L, N)$$

or the normalised form:

$$Y(M, L; N) = \frac{H(M, L) + H(N)}{H(M, L, N)}$$

in registration.

Using a Third Source of Information

This concept can be illustrated with the following simple example. Consider the registration of two one-dimensional signals $m(x)$ where $x \in [0,1]$ and $n(y)$, where $y \in [-\infty, +\infty]$ shown in Figure 10.4. We shall consider the two registration transformations $T : y \mapsto x$, denoted by T_R corresponding to correct alignment and T_E corresponding to a poor estimate. These are illustrated by the lower traces of Figure 10.4.

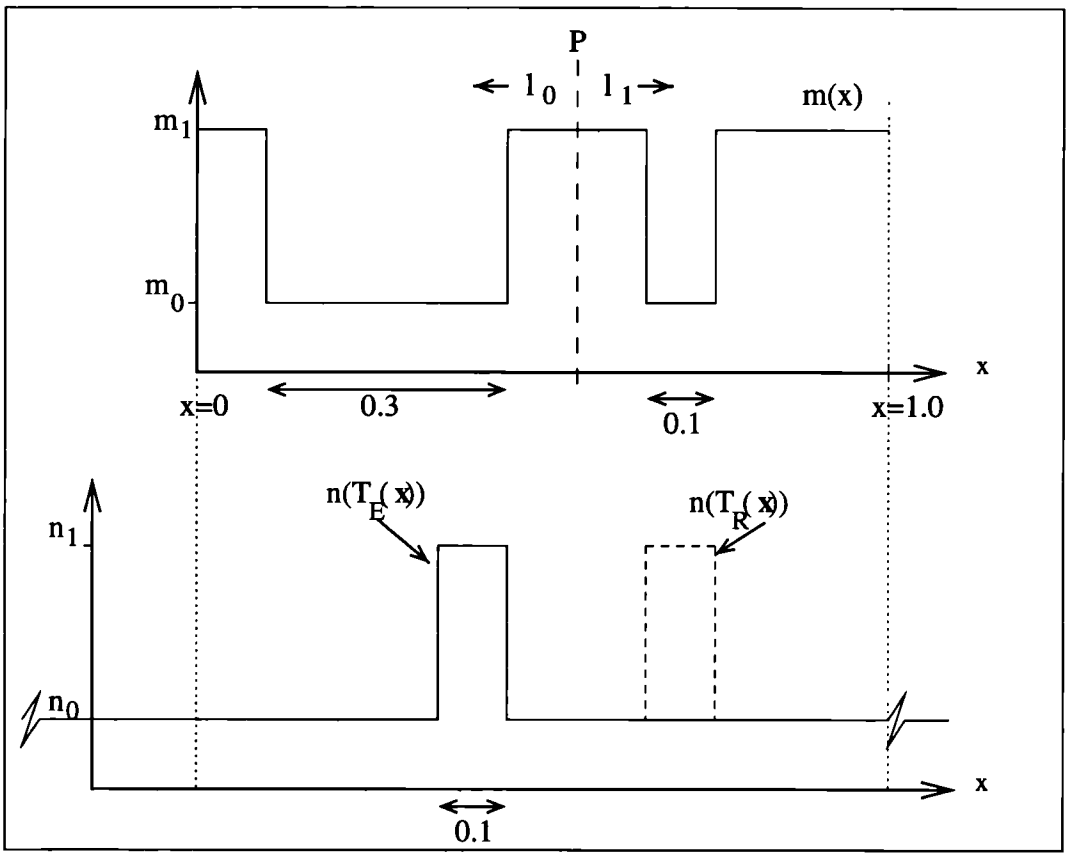


Figure 10.4: The registration of one dimensional signal $m(x)$ to $n(x)$ with additional labelling of M into $L\{l_0, l_1\}$

We can evaluate the joint probability distribution,

$$P(M,N) = \begin{bmatrix} p\{m_0,n_0\} & p\{m_1,n_0\} \\ p\{m_0,n_1\} & p\{m_1,n_1\} \end{bmatrix}. \tag{10.1}$$

over their region of overlap $T(y) \in [0,1]$, directly from the diagram so that,

$$P_R(M,N) = \begin{bmatrix} 0.3 & 0.6 \\ 0.1 & 0.0 \end{bmatrix}, \tag{10.2}$$

and,

$$P_E(M,N) = \begin{bmatrix} 0.3 & 0.6 \\ 0.1 & 0.0 \end{bmatrix}. \tag{10.3}$$

The task is to register the two functions using entropy measures of alignment. Here for both the alignments, the marginal entropies will remain the same so we need only consider the joint entropy,

$$\begin{aligned} H(M,N) = &-p\{m_0,n_0\}\log(p\{m_0,n_0\}) - p\{m_1,n_0\}\log(p\{m_1,n_0\}) \\ &-p\{m_0,n_1\}\log(p\{m_0,n_1\}) - p\{m_1,n_1\}\log(p\{m_1,n_1\}). \end{aligned} \tag{10.4}$$

Since the probability distribution for the two estimates $P_E(M,N) = P_R(M,N)$, the joint entropy will not distinguish between the two estimates,

$$H_E(M,N) = H_R(M,N). \tag{10.5}$$

Now, if we consider the addition of a third channel containing the label $l \in L$ where in this case $L = \{l_0, l_1\}$, we can now evaluate a three dimensional probability distribution which we can write out as,

$$P(M,N,L) = \begin{bmatrix} p\{m_0,n_0,l_0\}, & p\{m_1,n_0,l_0\}, & p\{m_0,n_0,l_1\}, & p\{m_1,n_0,l_1\} \\ p\{m_0,n_1,l_0\}, & p\{m_1,n_1,l_0\}, & p\{m_0,n_1,l_1\}, & p\{m_1,n_1,l_1\} \end{bmatrix}. \tag{10.6}$$

For the two transformation estimates this evaluates to,

$$P_R(M,N,L) = \begin{bmatrix} 0.3, & 0.2, & 0.0, & 0.4, \\ 0.0, & 0.0, & 0.1, & 0.0, \end{bmatrix}, \tag{10.7}$$

and,

$$P_E(M,N,L) = \begin{bmatrix} 0.2, & 0.2, & 0.1, & 0.4, \\ 0.0, & 0.1, & 0.0, & 0.0, \end{bmatrix}. \tag{10.8}$$

The joint entropies therefore evaluate to,

$$H_R(M,N,L) = 0.56$$

and,

$$H_E(M,N,L) = 0.74$$

(using $\log_{10}()$) allowing us to distinguish between the two estimates and align the images by minimising the joint entropy. In more typical cases where there is a varying area of overlap, we can maximise the normalised mutual information $Y(M,L;N)$.

10.3.3 Information and Region Labelling

In order to illustrate how extra information is being introduced by the labelling we can consider a simple region extraction process. This may consist of a thresholding to distinguish boundaries of interest, followed by a connected component labelling of the identified regions as illustrated in Figure 10.5. Thresholding is effectively combining intensities (symbols) in the image, producing only a subset of the information in the original image. If a binary threshold is used then the image is being reduced to two symbols $V = \{v_0, v_1\}$ where, if there are i intensities in the original image and m_v is the threshold intensity so,

$$v_0 = \{m_0, \ldots m_v\} \tag{10.9}$$

$$v_1 = \{m_{v+1}, \ldots m_i\}. \tag{10.10}$$

The information contained in this thresholded image is then given by,

$$H(V) = -\sum_{v \in \{v_0, v_1\}} p\{v\} \log p\{v\}. \tag{10.11}$$

Using the additivity property of information measures discussed in section 4.2.1, it can be shown that the information content of a set of symbols cannot be decreased when symbols are partitioned so,

$$H(m_0, \ldots m_v, m_{v+1} \ldots m_i) \geq H(v_0, m_{v+1}, \ldots m_i) \tag{10.12}$$

From this we can say that the information content of the thresholded image is always less than or equal to that in the original image,

$$H(V) \leq H(M). \tag{10.13}$$

In addition, the joint entropy between the image and its thresholded version is given by,

$$H(M,V) = -\sum_{v \in V} \sum_{m \in M} p\{m, v\} \log p\{m, v\} \tag{10.14}$$

which can be re-written in terms of the occurrence of the original intensities only,

$$\begin{aligned} H(M,V) = &-\sum_{m \in v_0} p\{m\} \log p\{m\} \\ &-\sum_{m \in v_1} p\{m\} \log p\{m\}, \end{aligned} \tag{10.15}$$

giving $H(M,V) = H(M)$. Using this, the mutual information between the image and its threshold,

$$I(M;V) = H(V) + H(M) - H(M,V). \tag{10.16}$$

becomes simply $I(M;V) = H(V)$, i.e. the information in the thresholded image $H(V)$ is simply a subset of that provided by the original image $H(M)$.

Labelling unconnected regions differently, effectively re-partitions the threshold labels and conversely can only maintain or increase the number of symbols (intensities or labels) present in our description of the scene. Applying the additivity property again, this must always maintain or increase the amount of information so that, if the threshold label v_0 is being divided up into unconnected regions $\{l_0 \dots l_j\}$,

$$H(l_0 \dots l_k, v_1) \geq H(v_0, v_1). \tag{10.17}$$

If $\{l_0, \dots l_j\}$ is simply a repartitioning of $\{m_0, \dots m_v\}$ (i.e. no regions with the same intensity are spatially unconnected) then $H(M,L) = H(M)$ and so $I(M,L;N) = I(M;N)$. If though, the image contains cases where an intensity $m \in v_0$ is partitioned into more than one label l then,

$$H(M,L) \geq H(M) \tag{10.18}$$

and similarly $H(M,L,N) \geq H(M,N)$ so,

$$I(M,L;N) \geq I(M;N). \tag{10.19}$$

By introducing connected region labelling information we can only maintain or increase the mutual information in the registration process. What we would like is a labelling for which $I(M;L)$ is minimised (there are no duplicated features in the labelling and the image) and $I(L;N)$ is maximised (there is optimum additional shared information with the other modality).

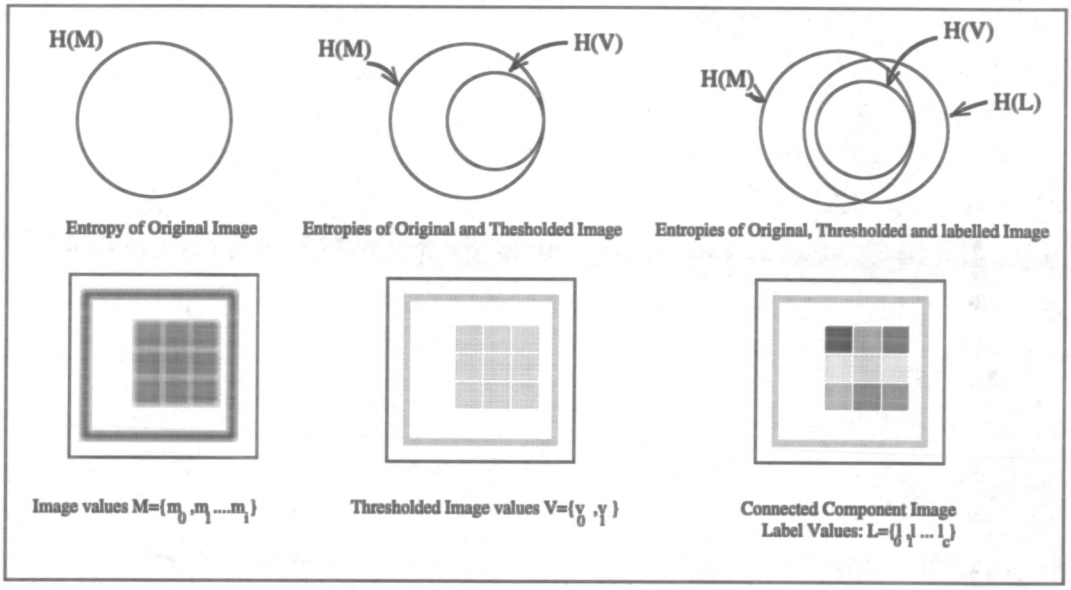


Figure 10.5: An illustration of the entropies of an image (left) where regions are blurred, its threshold (centre), and a connected component labelling of its threshold (right), with random gray levels to represent labels. Thresholding provides simply a subset of the original information, while a labelling which distinguishes between regions with the same value in the original image introduces additional information.

10.4 The Application of Region Labelling to MR-PET Pelvis Image Alignment

As discussed in chapter 9, it is common for both MR and PET ^{18}FDG scans of a patient to be acquired for the detection and staging of cervical cancer. Previously we have examined the problem of registration of this type of image data in the presence of severe intensity distortion which may occur in clinical MR scans of the pelvis. In this section we will investigate a second problem which can occur with these types of images, where connected region labelling may be useful. As with the previous work, the problem lies with the MR image, but in this case it is the underlying structure delineated by the MR image, rather than the artefacts introduced in image acquisition, which cause voxel similarity approaches to fail.

As described in chapter 9, we are interested in using the alignment of the bone structures delineated by the Fluoride tracer in PET, with the marrow and bone regions in the MR to constrain the alignment. It was observed that in some MR images, depending on the MR parameters and underlying anatomy, for marrow within the bone to exhibit the same MR values as regions of fat elsewhere around the pelvic region and abdomen.

10.4.1 Method

Region Labelling

The MR image was labelled interactively into four categories:

- Air
- Fat
- Bone Marrow
- Non-Fatty Intra-abdominal tissue

This was carried out using an interactive intensity based region growing algorithm to produce a third image with voxels set to one of four values. Figure 10.6 illustrates a slice from the labelled volume.

Evaluation of $p\{m, l, n\}$

The MR image was first re-sampled at the same resolution as the PET image. Linear interpolation was used to increase axial sampling. A Gaussian kernel was used to low pass filter the MR to give it a similar spatial resolution to that of the PET. For a given rigid transformation T all voxels in the region of overlap of the two images were used to

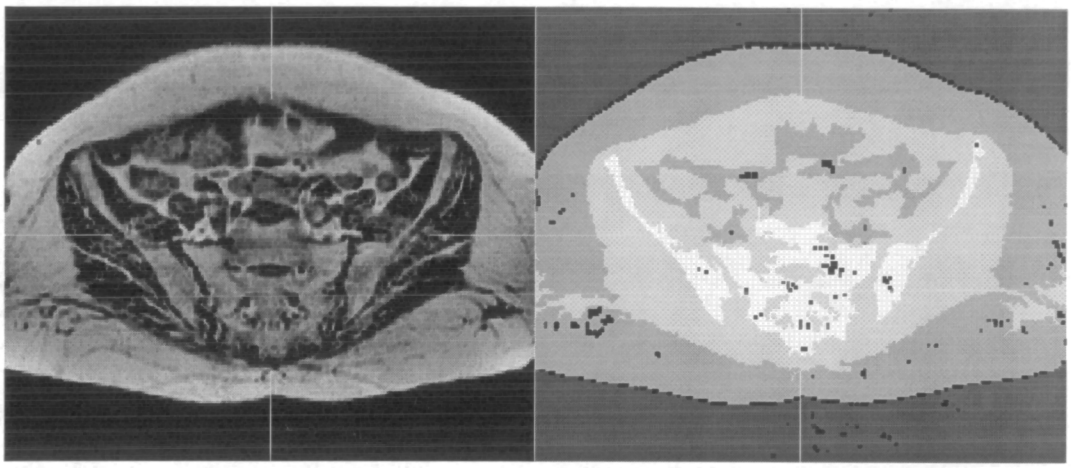


Figure 10.6: Example slice from MR volume (left) with labels (right).

evaluate $p\{m, l, n\}$. The MR was used as the reference image ($m(\mathbf{x})$) and corresponding PET ($n(\mathbf{y})$) intensity was estimated using tri-linear intensity interpolation. MR and PET intensities were binned into 64 levels each and four bins were used for the labels in L to form a discrete distribution. Estimates of $I(M;N)$ and $I(M,L;N)$ were then evaluated from this distribution.

10.4.2 Results

Figure 10.7 shows a plot of z axis displacement from the manual estimate of mutual information derived from only intensity, and of mutual information including labelling. The lower plot of $I(M;N)$ shows both a poor response to misalignment and an optimum appreciably displaced from the manual estimate. The addition of labelling into the measure $I(M,L;N)$ increases the overall mutual information for all displacements and produces a distinct optimum close to the manual estimate.

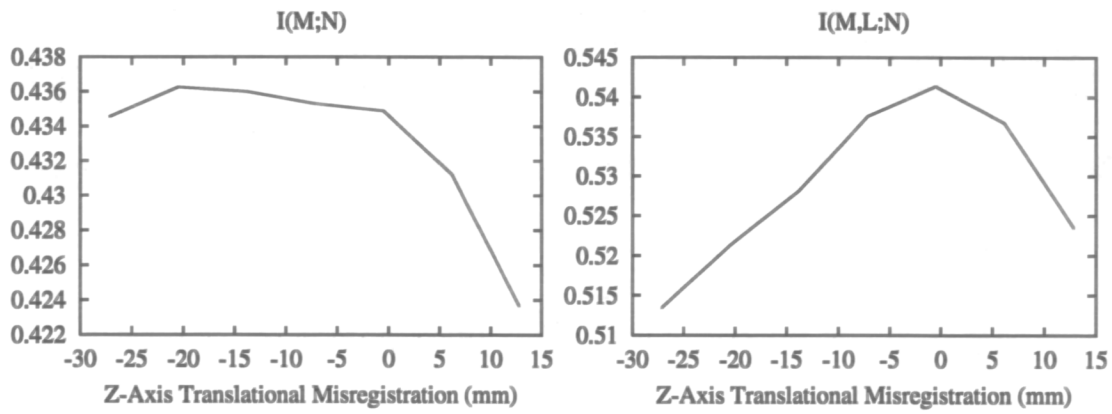


Figure 10.7: Graph of $I(M;N)$ and $I(M,L;N)$ with respect to axial (z) translation ($z = 0mm$ is the manual estimate).

A multi-resolution optimisation of $I(M, L; N)$ was carried out from a starting orientation defined by the scanner acquisitions, with the centres of the MR and PET volume aligned. The final transformation estimates are shown in table 10.1 along with those for optimisation of $I(M; N)$ and the manual estimate. The axes are as follows; x is from patient left to right, y is from patient front to back and z is from patient top to bottom. Registration by evaluation and optimisation of $I(M; N)$ derived from intensity only resulted in a visually poor registration. Multi-start of the optimisation and initialisation of the optimisation from the manual estimate confirmed that $I(M; N)$ was providing an incorrect global optimum corresponding to the alignment of bone features in PET with intra-abdominal fat. This resulted in both a large z axis displacement and a rotation around the y axis.

The difference between the manual estimate and the final automated estimate using the labelling approach $I(M, L; N)$ is within the expected accuracy of the manual registration procedure given the resolution of the original images. Interactive visual inspection of the results using colour overlay of PET intensity onto grey level MR in three orthogonal planes confirmed that the two results were visually comparable. Further work needs to be done in assessing final accuracy.

Method	Translation (mm)			Rotation (deg.)		
	t_x	t_y	t_z	θ_x	θ_y	θ_z
Manual	-6.0	34.5	39.0	2.5	0.0	-2.0
$I(M; N)$	-4.3	34.3	15.6	2.4	-3.0	-0.2
$I(M, L; N)$	-4.9	34.1	37.1	3.9	0.0	0.0

Table 10.1: Rigid transformation parameters estimated for the test image pair.

10.5 Summary

In this test example voxels were interactively labelled to provide additional information in the registration measure. By using an information theoretic approach which includes these labels the segmentation need not be ideal allowing the possibility of using an automated labelling scheme. Segmentation schemes distinguishing strongly between unconnected regions are favoured. Work is underway in looking at how such schemes can be integrated into the registration process. In particular our current multi-resolution approach to optimising mutual information could be extended to include a multi-resolution region extraction stage.

Further work is required to assess the extent to which labelling can be corrupted before the the final accuracy (location of the global optimum) is affected. The effect of typical errors produced by automated labelling schemes on the registration measure for a range of clinical images needs to be determined.

In the simple 1D example and on the clinical test image, labelling information differentiated regions which exhibited the same value in one modality but which exhibit different values in the other modality. Given a higher level knowledge about the regions it may be possible to use the same approach to combine regions of different intensity. This may be useful in reducing the influence of certain tissue boundaries which may have deformed between acquisitions.

This chapter has introduced a new type of voxel similarity approach that makes use of voxel label information as well as the original image intensities. Coarse segmentation of the higher resolution image is necessary to provide the label information. Unlike a surface matching approach [52], this technique does not require identification of corresponding boundaries from both modalities. Final alignment accuracy does not dependent on the accuracy of region labelling.

A possible interpretation of this technique is that the original intensity information is used for the accurate registration, and the labelling provides coarse scale information that removes unwanted optima far from the correct solution.

We have demonstrated this technique for the registration of MR and PET images of the pelvis, which could not be registered using conventional intensity based entropy measures. The registration solution was close to that produced using interactively identified point landmarks and produced a visually acceptable result.

Chapter 11

Conclusions

11.1 Summary of Findings

This thesis has investigated and developed one particular approach to the problem of automated retrospective 3D multi-modality image alignment. The approach chosen was to derive measures of image similarity which can be used as measures of image alignment. Such measures essentially summarise how well the occurrence of values in one modality explain the occurrence of values in the second modality. They provide a measure of alignment without the need to extract corresponding anatomical features.

11.1.1 Similarity Measures

The concepts behind measures of alignment were developed from a statistical description of the intersection or overlap of regions of material delineated by the two modalities. From this starting point alternative statistical measures were reviewed. Entropy and information theory was introduced as an approach to expressing alignment. The issue of normalisation of information measures with respect to changing image overlap was identified as an important factor.

11.1.2 Implementation

The task of evaluating 3D medical image similarity was examined in detail. Accounting for differences in spatial and sampling resolution, and excluding regions of padding used to store non-rectangular images can be essential in correctly evaluating similarity. A generic multi-resolution similarity evaluation and optimisation scheme was developed and used as a tool to compare the behaviour of both simple statistical and entropy based measures.

11.1.3 Comparison of Similarity Measures

In chapter 6 experiments illustrated that similarity measures, particularly normalised mutual information, can provide a robust and accurate measure of brain image alignment. The normalisation of entropy measures proved important in providing a behaviour robust to image field of view. The final algorithm based on the multi-resolution optimisation of normalised mutual information was equally applicable to both MR-PET and MR-CT alignment. For MR-CT alignment over 18 image pairs, this provided a maximum MIPVD of $2.5mm$ and a minimum MIPVD of $0.4mm$ from the bone implanted fiducial marker based estimate. For MR-PET alignment over 14 image pairs, this provided a maximum MIPVD of $6.5mm$ and a minimum MIPVD of $1.4mm$ from the fiducial marker based estimate. Many of the estimates were within the expected accuracy of the marker estimate. In one case of MR-PET alignment, visual inspection indicated errors in the fiducial marker estimate which were not present in the automated estimate.

11.1.4 Automated Brain Image Registration

In chapters 7 and 8 the behaviour of normalised mutual information coupled with multi-resolution optimisation was evaluated for use as a clinical brain image registration tool. This was with the aim of replacing conventional manual point based alignment which has been used at UMDS for more than 6 years. The system was applied to the alignment of a large database of clinical data including examples of many typical geometric and intensity artefacts. The rigid alignment of 25 out of 26 MR and CT, and all 14 MR and PET image pairs were recovered fully automatically. The estimates provided an accuracy visually comparable or superior to that of manual registration.

By extending the registration optimisation scheme to include estimates of orthogonal scaling and gantry tilt parameters, it is possible in a number of cases, to visually improve registration of clinical MR and CT. Here we are essentially correcting for limitations in the geometric calibration of the imaging modalities. Experimentation indicated that in cases of smaller axial field of view the estimation of extra parameters may decrease the precision of the overall transformation estimate. The retrospective recovery of scaling and skew parameters does though often provide a valuable check of the image geometry information.

11.1.5 Extension of Voxel Similarity Techniques

Chapters 9 and 10 examined approaches to extending the voxel similarity approach to the alignment of image pairs where the most probable measurement values do not correspond. Both chapters made use of information theory to introduce additional in-

formation into the measure of alignment. Chapter 9 dealt with the case where entropy measures fail because of distortion of measurement values across the imaged space. Partitioning of the imaged space and the evaluation of the measure across these partitions allow a fully automated registration with other modalities where conventional similarity measures fail.

Chapter 10 examined in more detail the limitations of similarity based alignment due to the nature of the underlying scene being imaged. Although voxel similarity cannot be used in all cases we can, for some images introduce topological information distinguishing unconnected regions of material exhibiting the same value. Using this approach it was demonstrated that it was possible to align MR and PET images of the pelvis where conventional entropy based measures provided an incorrect indication of alignment.

11.2 The Clinical Significance

11.2.1 Robustness and Automated Registration

A robust registration algorithm is one which can provide a clinically acceptable registration estimate between any image pair without the need for user intervention. The results presented here have been confirmed by clinical use of versions of algorithms developed in this thesis. Over the past two years this has been shown to provide fully automated registration in a number of applications. Clinical installations of software for MR-PET and MR-SPECT brain image alignment where this is the case include UMDS Guy's and St Thomas', the Hammersmith Hospital and Dartmouth College, New Hampshire, U.S.A.

In brain MR and CT alignment, unlike many publications in the field, this thesis has dealt with the alignment of non-stereotactic CT with MR. It has therefore tried to address the 'main market' for retrospective MR-CT alignment. Here, because of the limited X-ray dose, there is much smaller axial extent in CT and the registration problem is much greater. In the experiments in this thesis, the alignment of all but those images with the smallest extent (27mm in CT and 21mm in MR) were recovered fully automatically. This, in practice means that the majority of CT scans carried out for skull base surgery at our site may be aligned with MR automatically.

11.2.2 Accuracy and Uncertainty

Estimating a correspondence to an arbitrary accuracy between points in clinical image data is not possible. As discussed in section 6.3.1, it is very difficult even to define an accurate 'ground truth' transformation in a controlled imaging experiment. Although the method of interpolation used in evaluating a transformation estimate can be important

(in these experiments only linear interpolation was used), the three main factors really limiting the accuracy of registration of clinical images are:

- Geometric distortion of the image acquisition.
- Motion of the patient during image acquisition.
- Deformation of patient structures due to surgery, growth or disease progression or regression between acquisitions.

In essence, errors due to all of these can be treated as simply a limitation imposed by our choice of transformation to describe the spatial relationship between the two scans. The problem we have is that a transformation may differ locally, however shared structure which constrains alignment is not in general evenly distributed across the imaged space. There will be regions therefore where we will not be able to retrospectively evaluate local misalignment.

In registering images from the same modality, regions where we cannot determine alignment will commonly contain no structure in both images. Our uncertainty in the transformation in these regions may therefore not be clinically important. In the multi-modality case these regions of uncertain registration may contain important but different structures, for example blood vessels in one and tumour in another. In this case the uncertainty in our knowledge about the spatial relationship between these may be clinically critical. However we choose to model this transformation using any number of anatomical constraints, the fact remains that we can only ever infer with some level of uncertainty the local relationship between structures.

Many clinical images acquired for diagnosis contain artefacts, but these are visually evident to the trained radiologist who can take them into account in using them to make a diagnosis. When artefact might effect interpretation scans are repeated. On the whole, the effect of these limitations do not have a great impact on the bulk of clinical diagnosis. Similarly, in a registration facility, as long as there is some visual indication of a misalignment available to the clinician, then in many cases such a limitation may not be significant. It is better that the limitation is apparent than hidden or badly corrected.

Evaluating and presenting an estimate of the uncertainty of a registration may be equally as important as improving the registration itself. Ensuring a clinician is aware of registration accuracy in critical clinical situations requires both the development of new computational approaches and the training of radiologists and surgeons.

11.3 Further Work

11.3.1 Extension of Transformation Complexity

There have been many recent papers on deriving more complex geometric transformations between images from the same modality [23, 25] and also from different modalities [55]. There have also been some approaches to constraining alignment estimates using knowledge of tissue behaviour or shape [24, 29, 59]. We have seen in the work of this thesis that it is possible to include scaling and skew estimates in the registration to account for limitations in the scanner setup. We have also seen that when applied to a range of clinical data, this can effect the precision with which we can estimate registration parameters. It is possible to extend the approach to correct for patient deformation, more complex geometric distortion and patient motion which limit the accuracy of our registration estimate. There are though many transformations which make one image look like another and so we must treat these estimates with care.

The best approach must be to constrain the transformations with additional information we have about the images and scene. For deformation this means introducing information about the expected behaviour of different tissue types. For geometric distortion, this means making use of knowledge about how an image may be distorted (for example by the field strength in MR). Finally, for patient motion this means making use of knowledge of how a scan is acquired through time, and the types of patient motion which occur most often. All of these approaches mean that the technique becomes much less retrospective because, for example, we may need detailed knowledge about an acquisition sequence, not just the image data. Secondly it may make the technique much less automated where, for example, some user specification of tissue boundaries may be required.

11.3.2 Similarity Measures

The first two thirds of this thesis was concerned with the development of a measure derived solely from the co-occurrence of values in the two images. Essentially this made use of the assumption that the co-occurrence of the most probable values was maximised at registration. We can classify these as Type I measures, which do not require any further knowledge about the scene. Chapters 9 and 10 began to examine approaches to introducing additional information into the alignment process which we can term Type II.

This was achieved by essentially considering regions of local image similarity. In effect, by evaluating a measure over separate partitions we are allowing for 'special cases' of values being similar. In one area of image we are expecting value m_1 to correspond

to value n_1, while in another area we are expecting m_2 to correspond to n_1. In effect we are determining correspondence of values (similarity) by spatial context.

The main task in extending the application of type II measures is one of devising methods of partitioning the image space automatically, as was described for intensity distortion in chapter 9. Considerable experimentation is also required to validate the accuracy of estimates provided by these approaches.

11.3.3 Increasing the Speed of Registration

The simple multi-resolution evaluation and optimisation scheme has proved both robust and efficient in its use throughout the experiments in the thesis. The computational complexity and therefore time taken to recover alignment is dominated by the interpolation of corresponding measurements when forming a joint histogram. To provide high accuracy in the final estimate we must evaluate and optimise the registration estimate at high image sampling resolution. Currently, to recover a registration with acceptable accuracy requires between 15 and 45 minutes depending on the field of view of the images. In most cases, given the time to acquire and reconstruct the images, and transfer them to a common workstation, this time delay is not a major factor in clinical use.

In some cases though, for example in interactively examining a database of images, or when increasing the number of transformation parameters estimated, it may be useful to increase the speed of registration. There are two fundamentally different ways of reducing the computational cost of registration:

- Reducing the number of evaluations of a registration estimate by introducing more assumptions about the form of the parameter space into the optimisation scheme.
- Reducing the number of points in the image used to derive a measure of alignment.

There is a significant body of literature devoted to multi-dimensional optimisation [36] which may provide significant reductions in the number of evaluations without loss of robustness. However, the one which perhaps promises the greatest reduction in computation is that based on the selection of a subset of points with which to evaluate alignment.

A direct method of developing this technique is to look at the refinement of alignment at increasing image resolutions. Here we can extract an estimate with some given accuracy at a particular image resolution. When we increase the image resolution to improve the estimate we do not want to have to use all the higher resolution samples to evaluate alignment. What we want is a measure indicating which points in the higher resolution image will contribute most to improving our estimate of alignment.

11.3.4 Intra-operative Imaging

The ability to relate features delineated by high resolution pre-operative diagnostic images to points in the patient during surgery would be a major step in increasing the application of multi-modality diagnostic imaging.

Relating 3D Images to the 2D Surgical Scene

Registration of points from pre-operative 3D images to the 2D intra-operative scene is not a trivial problem. There have been many recent approaches to the task [19, 45], but none as yet provide a accurate, generally applicable solution. Essentially we must derive our estimate of alignment from one or more 2D views of a scene containing anatomical surfaces. Most approaches so far have dealt with the case where a single known surface is visible in the 2D view of the patient (e.g. bone or skin). In general though, for many surgical applications, as tissue is removed or deformed, we do not know which of the many anatomical surfaces within a 3D pre-operative image are visible in the 2D surgical scene. In estimating alignment we must also therefore search for the subset of boundaries delineated in the 3D modality which we can see in the 2D modality. Solving this problem robustly and automatically would be a significant step forward in surgery guidance.

Relating Diagnostic Images to Interventional Images

An alternative approach is to use interventional MR or CT imaging for guidance during surgery. In these systems it is generally not possible to produce the same high quality diagnostic images which are provided by conventional imaging systems. An alternative approach here is to use image similarity techniques to align high quality pre-operative images to the interventional images during surgery. The experiments on the alignment accuracy following the addition of noise in chapters 7 and 8, indicate that acceptable alignment estimates may be derived in cases of high noise levels. This suggests a possible approach to intra-operative imaging where very low quality images are acquired regularly and registered to high quality diagnostic images. Such an approach would permit the acquisition of, for example, very low dose X-ray CT images during a surgical procedure.

Bibliography

[1] P.E. Anuta. Spatial registration of multispectral and multitemporal digital imagery using fast fourier transform techniques. *IEEE Transactions on Geoscience Electronics*, 8:353–368, 1970.

[2] A. Apicella, J.H. Nagel, and R. Dura. Fast multi-modality image matching. In *Proc. Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, volume 10, pages 414–415. IEEE Comp. Soc. Press, Los Alamitos, CA., 1988.

[3] K.S. Arun, T.S. Huang, and S.D. Blostein. Least squares fitting of two 3-D point sets. *IEEE Transactions of Pattern Analysis and Machine Intelligence*, 9(5):698–703, 1987.

[4] L. Axel, J. Costantini, and J. Listerud. Intensity correction in surface-coil MR imaging. *AJR*, 148:418–420, 1987.

[5] S. L. Bacharach. The physics of positron emission tomography. In S.R. Bergmann and B.E. Sobel, editors, *Positron Emission Tomography of the Heart*, chapter 2, pages 13 43. Futura, 1992.

[6] S. Banerjee, D.P. Mukherjee, and D. Dutta Majumdar. Point landmarks for registration of CT and MR images. *Pattern Recognition Letters*, 16:1033–1042, 1995.

[7] H.G. Barrow, J.M. Tenenbaum, R.C. Bolles, and H.C. Wolf. Parametric correspondence and chamfer matching:two new techniques for image matching. In *Proc. 5th Int. Joint Conf. Artificial Intelligence*, pages 659–663, 1977.

[8] P.J. Besl and N.D. McKay. A method for registration of 3D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:239–256, 1992.

[9] J. Bezdek, L. Hall, and L. Clarke. Review of MR image segmentation techniques using pattern recognition. *Medical Physics*, 20:1033 1048, 1993.

[10] R. Boesecke, T. Bruckner, and G. Ende. Landmark based correlation of medical images. *Phys. Med. Biol.*, pages 121 126, 1989.

[11] G. Borgefors. An improved version of the chamfer matching algorithm. In *Proceedings of the 7th International Conference on Pattern Recognition*, volume 2, pages 1175 1177, 1984. Montreal, Canada.

[12] R.N. Bracewell. *The Fourier Transform and its Applications.* McGraw-Hill, New York, U.S.A., 1965.

[13] W. W. Brey and P.A. Narayana. Correction for intenisty falloff in surface coil magnetic resonance imaging. *Medical Physics*, 15:241 245, 1988.

[14] R.A. Brooks and G. Di Chiro. Principles of computer assisted tomography. *Phys. Med. Biol.*, 21(5):689 732, 1976.

[15] L. G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325 376, 1992.

[16] H. Chang and J.M. Fitzpatrick. A technique for accurate MR imaging in the presence of field inhomogeneities. *IEEE Trans. Med. Imaging*, 11:319 329, 1992.

[17] Y. Chen and G. Medioni. Object modelling by registration of multiple range images. *Image and Vision Computing*, 10:145 155, 1992.

[18] A. Colchester, P. Summers, and J. Zhao. Local MR angiography for neurosurgical guidance. In *Proceedings of the Society of Magnetic Resonance*, page 1169, 1995.

[19] A.C.F. Colchester, J. Zhao, K.S. Holton-Tainter, C.J. Henri, N. Maitland, P.T.E. Roberts, C.G. Harris, and R.J. Evans. Development and preliminary evaluation of VISLAN, a surgical planning and guidance system using intra-operative video imaging. *Medical Image Analysis*, 1(1):73–90, 1996.

[20] A. Collignon, D. Vandermeulen, P. Suetens, and G. Marchal. Registration of 3D multi-modality medical images using surfaces and point landmarks. *Pattern Recognition Letters*, 15:461–467, 1994.

[21] A. Collignon, D. Vandermeulen, P. Suetens, and G. Marchal. 3D multi-modality medical image registration using feature space clustering. In *Proceedings of the 1st International Conference on Computer Vision, Virtual Reality and Robotics in Medicine*, pages 195–204. Spinger-Verlag, 1995. Nice.

[22] A. Collignon, D. Vandermeulen, P. Suetens, G. Marchal, A. Baert, and A. Oosterlinck. Automatic registration of 3D images of the brain based on fuzzy objects. In *Proceedings of Medical Imaging 1994*, volume 2167, page N.A. SPIE Press, 1994. Newport Beach, California.

[23] D.L. Collins, A.C. Evans, C. Holmes, and T.M. Peters. Automatic 3D segmentation of neuro-anatomical structures from MRI. In Bizais Y., Barillot C., and Di Paola R., editors, *Proceedings of Information Processing in Medical Imaging*, pages 139–152. Kluwer Academic Publishers, 1995. Brest, France.

[24] D.L. Collins, G.Le Goualher, R. Venugopal, A. Caramanos, A.C. Evans, and C. Barillot. Cortical constraints for non-linear cortical registration. In K.H. Hohne and R. Kikinis, editors, *Proceedings of the 4th International Conference on Visualization in Biomedical Computing*, pages 307–316. Springer-Verlag, 1996.

[25] G.E. Cristensen, M.I. Miller, and M. Vannier. A 3D deformable magnetic resonance textbook based on elasticity. In *Proceedings if AAAI Spring Symposium Series: Applications of Computer Vision in Medical Image Processing*, pages 153–156, 1994.

[26] P.A. Van den Elsen and M.A. Viergever. Marker guided multimodality matching of the brain. *European Radiology*, 4:45-51, 1994.

[27] A. P. Dhawan, L. K. Arata, A. V. Levy, and J. Mantil. Iterative prinicple axes registration method for analysis of MR-PET brain images. *IEEE Transactions on Biomedical Engineering*, 42(11):1079–1087, 1995.

[28] R. O Duda and P. E. Hart. *Pattern Calssification and Scene Analysis.* John Wiley, New York, 1973.

[29] P.J. Edwards, D.L.G. Hill, J.A. Little, and D.J. Hawkes. Deformation for image guided interventions using a three component tissue model. In J. Duncan and G. Gindi, editors, *Proceedings of 15th International Conference on Information Processings in Medical Imaging*, pages 218–231. Springer-Verlag, 1997.

[30] A.C. Evans, C. Beil, S. Marrett, C.J. Thompson, and A. Hakim. Anatomical-functional correlation using and adjustable MRI-based region of interest atlas with positron emission tomography. *Journal of Cerebral Blood Flow and Metabolism*, 8:513–530, 1988.

[31] A.C. Evans, S. Marrett, L. Collins, and T.M. Peters. Anatomical-functional correlative analysis of the human brain using three dimensional imaging systems. In R. Schneider, S. Dwyer III, and R. Jost, editors, *Medical Imaging: Image Processing*, volume 1092, pages 264–274. SPIE Press, 1989.

[32] A.C. Evans, S. Marrett, J. Torrescorzo, and L. Collins S. Ku. MRI-PET correlation in three dimensions using a volume-of-interest (voi) atlas. *Journal of Cereb. Blood Flow Metab.*, 11:A69–A78, 1991.

[33] T. L. Faber and E. M. Stokely. Orientation of 3-D structures in medical images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10:626 633, 1988.

[34] J. Feldmar, G. Malandain, J. Declerck, and N. Ayache. Extension of the icp algorithm to non-rigid intensity-based registration of 3D volumes. In *Proceedings*

of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis, pages 84–93. IEEE Computer Society Press, 1996.

[35] R.A. Fisher. Theory of statistical estimation. *Proceedings of the Cambridge Philosophical Society*, 22:700–725, 1925.

[36] R. Fletcher. *Practical Methods of Optimization.* John Wiley, Chichester, England, 1987.

[37] K.J. Friston, J. Ashburner, J.B. Poline, C.D. Frith, J.D. Heather, and R.S.J. Frackowiak. Spatial registration and normalisation of images. *Human Brain Maping*, 2:165–189, 1995.

[38] I. Gath and A. B. Geva. Unsupervised optimal fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Inte lligence*, 11(7):773 781, 1989.

[39] Y. Ge, M Fitzpatrick, J.R. Votaw, S. Gadamsetty, R.J. Maciunas, R.M. Kessler, and R.A. Margolin. Retrospective registration of PET and MR brain images: An algorithm and its stereotatic validation. *Journal of Computer Assisted Tomography*, 18:800–810, 1994.

[40] T. Geraud, I. Bloch, A. Collignon, H. Maitre, and P. Suetens. Matching criteria for 3-D multimodality medical image registration. Technical report, Katholieke Universiteit Leuven and Telecom Paris, December 1992.

[41] P. Gerlot-Chiron and Yves Bizais. Registration of multimodality medical images using a region overlap criterion. *Computer Vision, Graphics and Image Processing: Graphical Models and Image Processing*, 54:396–406, 1992.

[42] G.E. Goldberg. *Genetic Algorithms in Search optimisation and Machine Learning.* Addision Wesley, Mass. U.S.A., 1989.

[43] R.C. Gonzalez and R.E. Woods. *Digital Image Processing.* Addison Wesley, New York, 1992.

[44] A. Gueziec and N. Ayache. Smoothing and matching of 3D space curves. In *Proceedings of the Second European Conference on Computer Vision*, 1992. Santa Margherita Ligure, Italy.

[45] N. Hata, W.M. Wells, M. Halle, S. Nakajima, P. Violoa, R. Kikinis, and F.A. Jolesz. Image guided microscope surgery system using mutual-information based registration. In K.H. Hohne and R. Kikinis, editors, *Proceedings of the 4th International Conference on Visualization in Biomedical Computing*, pages 317–326. Springer-Verlag, 1996.

[46] D.L.G. Hill and D.J. Hawkes. Medical image registration using knowledge of adjacency of anatomical structures. *Image and Vision Computing*, 12:173 178, 1994.

[47] D.L.G. Hill, D.J. Hawkes, J.E. Crossman, M.J. Gleeson, T.C.S. Cox, E.E.C.M.L. Bracey, A.J. Strong, and P. Graves. Registration of MR and CT images for skull base surgery using point-like anatomical features. *British Journal of Radiology*, 64:1030–1035, 1991.

[48] D.L.G. Hill, D.J. Hawkes, N.A. Harrison, and C.F. Ruff. A strategy for automated multimodality image registration incorporating anatomical knowledge and imager characteristics. In *Proceedings of Information Processing in Medical Imaging*, pages 182 196, 1993. Flagstaff,U.S.A.

[49] D.L.G. Hill, C. Studholme, and D.J. Hawkes. Voxel similarity measures for automated image registration. In *Proceedings of Visualisation in Biomedical Computing*, volume SPIE 2359, pages 205–216. SPIE Press, 1994. Rochester Mn.,U.S.A.

[50] B.L. Holman, R.E. Zimmerman, K.A. Johnson, P.A. Carvalho, R.B. Schwartz, J.S. Loeffler, E. Alexander, C.A. Pelizzari, and G.T.Y Chen. Computer-assisted superimposition of magnetic resonance and high-resolution technetium-99m-HMPAO and thallium-201 SPECT images of the brain. *J. Nucl. Med.*, 32(8):1478–1484, 1991.

[51] J.J. Jacq and C. Roux. Registration of 3-D images by genetic optimisation. *Pattern Recognition Letters*, 16:823–841, 1995.

[52] H. Jiang, R.A. Robb, and K.S. Holton. New approach to 3-D registration of multimodality medical images by surface matching. In *Proceedings of Visualisation in Biomedical Computing*, pages 196–213. SPIE Press, 1992. Chapel Hill, N.C.,U.S.A.

[53] I. Kapouleas, A Alavi, W.A. Alves, R.E. Gur, and D.W. Weiss. Registration of three-dimensional MR and PET images of the human brain without markers. *Radiology*, 181:731–739, 1991.

[54] P.J. Kelly. Compuer-assisted sterotaxis: New approaches for the management of intracranial intra-axial tumors. *Neurol*, 36:353–541, 1986.

[55] B. Kim, J. Boes, K.A. Frey, and C.R. Meyer. Mutual information for automated multimodal image warping. In K.H. Hohne and R. Kikinis, editors, *Proceedings of the 4th International Conference on Visualization in Biomedical Computing*, pages 349 354. Springer-Verlag, 1996.

[56] M.N.J. Knufman, P.A. Van den Elsen, J.P.M. Cillessen, J.W. Van Isselt, and C.A.F. Tulleken. Spatial integration of multimodal brain images in cerebral infarction. *Brain Topogr*, 5(2):165–170, 1992.

[57] S. Lavalle. Registration for computer integrated surgery. In R.H. Taylor, S. Lavallee, and G.C. Burdea, editors, *Computer Integrated Surgery*, pages 77–97. MIT Press, 1995.

[58] D.N. Levin, X Hu, K.K. Tan, S. Galhotra, C.A. Pelizzari, G.T.Y. Chen, R.N. Beck, C.T. Chen, M.D. Cooper, J.F. Mullan, J. Hekmatpanah, and J.P. Spire. The brain: Integrated three-dimensional display of MR and PET images. *Radiol.*, 175(3):783–789, 1989.

[59] J.A. Little, D.L.G. Hill, and D.J. Hawkes. Deformations incorporating rigid structures. *Computer Vision and Image Processing*, 66(2):223 232, 1997.

[60] R.B. Lufkin, T. Sharpless, B. Flannigan, and W. Hanafee. Dynamic-range compression in surface-coil MRI. *AJR*, 147:379 382, 1986.

[61] Y. Ma, M. Kamber, and A.C. Evans. 3D simulation of PET brain images using segmented MRI data and positron tomograph characteristics. *Computerised Medical Imaging and Graphics*, 17(4/5):354 371, 1993.

[62] A. Macovski. Noise in MRI. *Magnetic Resonance in Medicine*, 36:494–497, 1996.

[63] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximisation of mutual information. In *Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*, pages 14–22. IEEE Computer Society Press, 1996.

[64] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximisation of mutual information. *IEEE Transactions on Medical Images*, 16(2):187–198, 1997.

[65] J.B.A. Maintz, P.A. van den Elsen, and M.A. Viergever. Evaluation of ridge seeking operators for multimodality medical image matching. *IEEE Transactions on Medical Images*, 18(4):353–365, 1996.

[66] P. Mansfield and P.G. Morris. *NMR Imaging in Biomedicine.* Academic Press, Orlando, Florida, U.S.A., 1982.

[67] C.R. Maurer, G.B. Aboutanos, B.M. Dawant, R.J. Maciunas, and J.M. Fitzpatrick. Registration of 3-D images using wieghted geometrical features. *IEEE Transactions on Medical Imaging*, 15(6):1163–1178, 1996.

[68] C.R. Maurer and J.M. Fitzpatrick. A review of medical image registration. In R.J. Macuinas, editor, *Interactive Image-Guided Neurosurgery*, pages 14–44. Am. Assoc. Neurological Surgeons, 1993.

[69] C.R. Meyer, G.S. Leichtman, J.A. Brunberg, R.L. Wahl, and L.E. Quint. Simultaneous usage of homologous points, lines, and planes for optimal, 3-D, linear registration of multimodality imaging data. *IEEE Transactions on Medical Imaging*, 14:1 11, 1995.

[70] A. Mitiche and J.K. Aggawal. Contour registration by shape specific points for shape matching. *Computer Vision, Graphics and Image Processing*, 22:396–408, 1983.

[71] R.N. Nagel and A. Rosenfeld. Ordered search techniques in template matching. *Proceedings of the IEEE*, 60:242–244, 1972.

[72] M.A. Oghabian and A. Todd-Pokropek. Registration of brain images by multiresolution sequential method. In *Proceedings of Information Processing and Medical Imaging*, pages 165–174. Springer-Verlag, 1991. London.

[73] C.A. Pelizzari, G.T.Y. Chen, D.R. Spelbring, R.R. Weichselbaum, and C-T Chen. Accurate three dimensional registration of CT, PET and/or MR images of the brain. *Journal of Computer Assisted Tomography*, 13:20–26, 1989.

[74] M.J.D. Powell. An efficient method for finding the minimum of a function of several variables without calculating derivatives. *Comput. J.*, 7:155–163, 1964.

[75] A. Rangarajan, H. Chui, and F. Bookstein. The softassign procrustes matching algorithm. In J. Duncan and G. Gindi, editors, *Proceedings of 15th International Conference on Information Processings in Medical Imaging*, pages 29–42. Springer-Verlag, 1997.

[76] F.M. Reza. *An Introduction to Information Theory.* Dover, New York, 1994.

[77] J.T. Richtsmeier, C.H. Paik, P.C. Elfert, T.M. Cole, and H.R. Dahlman. Precision, repeatability and validation of the localisation of cranial landmarks using computed tomography. *Cleft Palate-Craniofacial Journal*, 32(3):218–227, 1995.

[78] R. Rohr, H.S. Stiehl, R. Sprengel, W. Beil, T.M. Buzug, J. Weese, and M.H. Kuhn. Point-based elastic registration of medical image data using approximating thin-plate splines. In K.H. Hoehne and R. Kikinis, editors, *Proceedings of Visualisation in Biomedical Computing*, pages 297 306. Spinger-Verlag, 1996. Hamburg.

[79] A. Rosenfeld and G.J. Vanderbrug. Coarse-fine template matching. *IEEE Transactions on Systems, Man and Cybernetics*, 7:104–107, 1977.

[80] C.F. Ruff, D.L.G. Hill, and D.J. Hawkes. A programmable user interface for multimodal medical image processing. In *Proceedings of IEEE-EMBS*, pages 155–160, 1992. Rennes.

[81] H. Rusinek, W.H. Tsui, A. V. Levy, M. E. Noz, and M. J. de Leon. Principle axes and surface fitting methods for three-dimensional image registration. *IEEE Transactions on Biomedical Engineering*, 34:2019–2024, 1993.

[82] L. Schad, R. Boesecke, W. Schlegel, G.H. Hartmann, V. Sturm, L.G. Strauss, and W. J. Lorenz. Three dimensional image correlation of CT, MR and PET studies in radiotherapy treatment planning of brain tumours. *Journal of Computer Assisted Tomography*, 11:948 954, 1987.

[83] A.M. Scott, H.A. Macaplinlac, C.R. Divgi, J.J. Zhang, H. Kalaigian, K. Pentlow, S. Hilton, M.C. Graham, G. Sgouros, C. Pelizzari, G. Chen, J. Schlom, S.J. Goldsmith, and S.M. Larson. Clinical validation of SPECT and CT/MRI image registration in radiolabeled monocloncal antibody studies of colorectal carcinoma. *Journal of Nuclear Medicine*, 35(12):1976–1984, 1994.

[84] J. Skilling. Fundamentals of maxent in data analysis. In B. Buck and V. A. Macaulay, editors, *Maximum Entropy in Action*, pages 19 40. Oxford Unviersity Press, 1991.

[85] C. Studholme, D.J. Hawkes, and D. L. G. Hill. Automated 3D registration of MR and CT images of the head. *Medical Image Analysis*, 1(2):163–175, 1996.

[86] C. Studholme, D.J. Hawkes, and D. L. G. Hill. Robust fully automated 3D registration of MR and PET images of the brain using multiresolution voxel similarity measures. *Medical Physics*, 24(1):25–35, 1997.

[87] C. Studholme, D.L.G. Hill, and D.J. Hawkes. Multiresolution voxel similarity measures for MR-PET registration. In Bizais Y., Barillot C., and Di Paola R., editors, *Proceedings of Information Processing in Medical Imaging*, pages 287–298. Kluwer Academic Publishers, 1995. Brest, France.

[88] M. Svedlow and C.D. McGillem. Image registration: Similarity measure and preprocessing method comparison. *IEEE Transactions on Aerospace and Electronic Systems*, 14:141 149, 1978.

[89] J.P. Thirion, A. Gourdon, O. Monga, A. Gueziec, and N. Ayache. Automatic registration of 3D CAT-scan images using surface curvature. In *3D Advanced Image Processing in Medicine: Proceedings of the 14th Annual Conference of the IEEE EMBS*, pages 85–90, 1992. Rennes.

[90] D.G.T Tomas, S.S. Gill, C.B. Wilson, J.L. Darling, and C.S. Parkins. Use of a relocatable stareotactic frame to integrate positron emission tomography and computed tomography images: Applications in human malignant brain tumours. *Stereotactic Functional Neurosurgery*, 54+55:338 392, 1990.

[91] T.G. Turkington, J.M. Hoffman, R.J. Jaszczak, J.R. MacFall, C.C. Craig Harris, C.D. Kilts, C.A. Pelizzari, and R. Edward Coleman. Accuracy of surface fit registration for PET and MR brain images using full and incomplete brain surfaces. *Journal of Computer Assisted Tomography*, 19:117–124, 1995.

[92] M. Unser, P. Thevanaz, and L. Yaroslavsky. Convolution-based interpolation for fast, high-quaility rotation of images. *IEEE Transactions on Image Processing*, 4:1371–1381, 1995.

[93] P.A. van den Elsen, J.B.A. Maintz, E.J.D. Pol, and M.A. Viergever. Automatic registration of CT and MR brain images using correlation of geometrical features. *IEEE Transactions on Medical Images*, 14(2):384–398, 1995.

[94] P.A. van den Elsen, J.B.A. Maintz, and M.A. Viergever. Geometry driven multimodality matching of brain images. *Brain Topography*, 5:153–158, 1992.

[95] P.A. van den Elsen, E.J.D. Pol, T.S. Sumanawaeera, P.F. Hemler, S. Napel, and J.R. Adler. Grey value correlation techniques used for automatic matching of CT and MR brain and spine images. In *Proceedings of Visualisation in Biomedical Computing*, pages 227–237. SPIE Press, 1994. Rochester Mn.,U.S.A.

[96] P.A. van den Elsen, E.J.D. Pol, and M.A. Viergever. Medical image matching- a review with classification. *IEEE Transactions in Medicine and Biology*, 12(1):26-39,1993.

[97] M. van Herk and H.M. Kooy. Automatic three-dimensional correlation of CT-CT, CT-MRI, and CT-SPECT using chamfer matching. *Medical Physics*, 21:1163–1178, 1994.

[98] K. Vincken. *Probabilistic Mutliscale Image Segmentation by the Hyperstack.* PhD thesis, Universiteit Utrecht, 1995.

[99] P.A. Viola and W.M. Wells. Alignment by maximisation of mutual information. In *Proceedings of the 5th International Conference on Computer Vision*, pages 15 23, 1995.

[100] S. Webb. *The Physics of Medical Imaging.* Institute of Physics, London, 1988.

[101] W. Wells, Grimson W, R. Kikinis, and F. Jolesz. Statistical intensity correction and segmentation of MRI data. In *Proceedings of Visualisation in Biomedical Computing*, volume SPIE 2359, pages 13–24. SPIE Press, 1994. Rochester, Minnesota, U.S.A.

[102] W.M. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis. Multi-modal volume registration by maximisation of mutual information. *Medical Image Analysis*, 1(1):35–51, 1996.

[103] J. West, J.M. Fitzpatrick, and et al. Comparison and evaluation of retrospective intermodality registration techniques. *Journal of Computer Assisted Tomography*, 21:554–566, 1997.

[104] J.C.H. Wong, C. Studholme, D.J. Hawkes, and M.N. Maisey. Evaluation of the limits of visual detection of image misregistration in a brain ^{18}FDG-MRI study. *European Journal of Nuclear Medicine*, 24:642 650, 1997.

[105] R.Y. Wong and E.L. Hall. Hierarchical scene matching. *IEEE Transactions on Computers*, 27(4):359–366, 1978.

[106] W.L. Wong, C. Studholme, P. Lewis, K.S. Raju, R.P. Beaney, K. Tonge, T. Nunan, D.J. Hawkes, and J. Pemberton. Combined MR,CT and PET imaging in oncological patients. *British Journal of Radiology*, 66(suppl):33 34, 1993.

[107] R.P. Woods, J.C. Mazziotta, and S.R. Cherry. A rapid automated algorithm for accurately aligning and reslicing PET images. *Journal of Computer Assisted Tomography*, 16:620–633, 1992.

[108] R.P. Woods, J.C. Mazziotta, and S.R. Cherry. MRI-PET registration with automated algorithm. *Journal of Computer Assisted Tomography*, 17:536–546, 1993.

[109] S.S. Yao and J.R. Gilbert. Registration of a synthetic aperture radar image to thematic mapper imagery for remote sensing applications. *IEEE Transactions on Geoscience and Remote Sensing*, 6:557–563, 1984.

Selected Publications

Articles in Refereed Journals

- C. Studholme, D.L.G.Hill,D.J. Hawkes,*Automated Three-dimensional Registration of Magnetic Resonance and Positron Emission Tomography Brain Images by Multiresolution Optimisation of Voxel Similarity Measures*, **Medical Physics**, Vol 24(1), pp 25-35, Jan 1997.

- J.C.H. Wong, C. Studholme, D.J. Hawkes, M.N. Maisey, *Evaluation of the Limits of Visual Detection of Image Misregistration in a Brain ^{18}FDG-MRI Study*, **European Journal of Nuclear Medicine**, Vol 24, pp. 642-650, 1997.

- C. Studholme, D.L.G.Hill, D.J. Hawkes, *Automated 3D Registration of MR and CT Images of the Head*, **Medical Image Analysis**, Vol. 1(2),pp 163-175, 1996.

- J. West, J.M. Fitzpatick, M.Y. Wang, B.M. Dawant, C.R. Maurer, R.M. Kessler, R.J. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P.A. van den Elsen, S. Naple,T.S. Sumanaweera, B. Harkness, P.F. Hemler, D.L.G. Hill, D.J. Hawkes, C. Studholme, J.B.A. Maintz, M.A. Viergever, G. Malandain, X. Pennec, M.E. Noz, G.Q. Maguire, M. Pollack, C.A. Pelizzari, R.A. Robb, D. Hanson, R.P. Woods, *Comparison and evaluation of retrospective intermodality image registration techniques*, **Journal of Computer Aided Tomography**, Vol 21, pp. 554-566, 1997.

- A.J. Gandhe, D.L.G. Hill, C. Studholme, D.J. Hawkes, C.F. Ruff, A.J. Strong, T.C.S. Cox, M.J. Gleeson,*Combined and 3D Rendered Multimodal Data for Planning Skull Base Surgery: A prospective Evaluation*, **Neurosurgery**, Vol 35(3), pp. 463-471, 1994.

- D.L.G. Hill, D.J. Hawkes, T.C.S. Cox, A.J. Strong, W.L. Wong, C.F. Ruff, N. Kitchen, D.G.T. Thomas, J.E. Crossman, C. Studholme, A.J. Gandhe, S.E.M. Green, G.P. Robinson, *Accutate Frameless Registration of MR and CT Images of the Head: Applications in Surgery and Radiotherapy Planning*, **Radiology**, 191, pp. 447-454, 1994.

Refereed Articles in Conference Proceedings

- C. Studholme, P.E. Summers, D.L.G. Hill, D.J. Hawkes, *An Information Theory Approach to the Alignment of Images Containing Measurement Inhomogeneity: Application to MR Surface Coil Angiography of the Brain*, Eds. C.J. Taylor, J.A. Noble, J.M. Brady, **Proceedings of Medical Image Understanding and Analysis 97**, pp. 45-48, July 1997.

- C. Studholme, D.L.G. Hill, D.J. Hawkes, *Auomated Multi Modality Registration using the Full Affine Transformation: Application to MR and CT Guided Skull Base Surgery*, Eds. K.H. Hohne, R. Kikinis, **Proceedings of Visualisation in Biomedical Computing**, Springer-Verlag, pp. 601-606, 1996.

- C. Studholme, D.L.G. Hill, J. Wong, D.J. Hawkes, M.N. Maisey, *Registration Measures for Automated 3D Alignment of PET and Intensity Distorted MR Images*, Eds. K. Mardia C.A. Gill, I.L. Dryden, **Proceedings of Image Fusion and Shape Variability Techniques**, Leeds University Press, pp. 186-193, 1996.

- C. Studholme, D.L.G. Hill, D.J. Hawkes, *Incorporating Connected Region Labelling into Automated Image Registration Using Mutual Information*, **Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis**, IEEE Computer Society Press, pp. 23-31, 1996.

- C. Studholme, D.L.G. Hill, D.J. Hawkes, *Automated 3D registration of truncated MR and CT images of the head*, Eds. D. Pycock, **Proceedings of the British Machine Vision Conference**, BMVA Press, 1995, pp. 27-36, 1995.

- C. Studholme, D.L.G. Hill, D.J. Hawkes, *Multiresolution voxel similarity measures for MR-PET registration*, Eds. Y. Bizais , C. Barillot, R. Di Paola, **Proceedings of Information Processing in Medical Imaging**, Kluwer Academic Publishers, pp. 287-298, 1995.